ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/200787272

Carbohydrates

Chapter · January 2008

CITATIONS
0

READS
27,170

3 authors:

Suman Khowala
47 PUBLICATIONS **391** CITATIONS
SEE PROFILE

Deepak Verma
Manovikas Kendra Rehabilitation and Research Institute for the Han...
29 PUBLICATIONS **111** CITATIONS
SEE PROFILE

Samudra Prosad Banik
University of Calcutta
20 PUBLICATIONS **282** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project
Mitochondrial involvement in the pathophysiology of neurodevelopmental disorders, ADHD and ASD View project

Project
Association of monoamine transporter genes with autism View project

All content following this page was uploaded by Deepak Verma on 22 May 2014.

The user has requested enhancement of the downloaded file.

BIOMOLECULES: (INTRODUCTION, STRUCTURE & FUNCTION)

Carbohydrates

Suman Khowala, Deepak Verma, Samudra P. Banik
Drug Developement and Biotechnology
Indian Institute of Chemical Biology
4, Raja S. C. Mullick Road.
Kolkata-700032. INDIA
Email: sumankhowala@iicb.res.in

Revised 04-June-2008

CONTENTS

- Introduction
- Definition
  - Importance of carbohydrates
  - Biosynthesis of other compounds
- Classification
- Isomerism
  - Chirality
  - Structural isomerism
  - Stereoisomerism
  - Mutarotation
  - Configuration of glucose
- Important reactions of Aldehydes and Ketones
- Important reactions of sugars
- Reactions of sugar due to hydroxyl groups
- Monosaccharide, Disaccharides, Trisaccharide and Polysaccharide
  - Monosaccharide
  - Disaccharide
  - Trisaccharides
  - Polysaccharides
    - Homopolysaccharides
    - Heteropolysaccharides
- Sugar derivatives
  - Derivatives of Monosaccharide
    - Amino sugars and N-acetylated sugars
    - Carboxylic acid sugars
    - Sugar phosphates
    - Sugar alcohols
  - Derivatives of polysaccharides
- Carbohydrates in living system
  - ABO Blood group system
  - Sugars in the cell wall of bacteria
  - Sugars in the cell wall of fungus
  - Derivatives with proteins
    - Glycoproteins (Glycophorins, Lectins, Mucins)

Proteoglycans
Conformation of polysaccharides
Three dimensional orientation
Orientation of the glycosidic groups
Rare sugars
Carbohydrate assays

Keywords
Carbohydrates; Sugars; Monosaccharides; Disaccharides; Trisachharides; Polysaccharides; Aldoses; Ketoses; Stereoisomerism; Optical isomerism; Mutarotation; Conformation of sugars; Sugar derivatives; Reactions of sugars.

Introduction

Carbohydrates are the most abundant biomolecules belonging to class of organic compounds found in living organisms on earth. Each year, more than 100 billion metric tons of CO_2 and H_2O are converted into cellulose and other plant products due to photosynthesis. Living matter is largely made of biomolecules consisting of water and complex polymers of amino acids, lipids, nucleotides and carbohydrates. Carbohydrates are most special of them in that they remain associated with the three other polymers mentioned. Carbohydrates are linked with amino acid polymers (proteins) forming glycoproteins and with lipids as glycolipids. Carbohydrates are present in DNA and RNA, which are essentially polymers of D-ribose-phosphate and 2-deoxy-D-ribose phosphate to which purines and pyrimidines bases are attached at the C-1 reducing position. Carbohydrates are a widely diverse group of compounds that are ubiquitous in nature. More than 75% of the dry weight of the plant world is carbohydrate in nature - particularly cellulose, hemicellulose and lignin.

Existence of sugars is confirmed before life itself appeared on earth. It is highly probable that there was a relative abundance of various sugars and their phosphates in the prebiotic world, where the basic building blocks must have polymerized and assembled, ultimately to form a self-sustaining, self-reproducing, and adaptive entity. The reality or possibility that carbohydrates will be found in fossils is almost zero because they are relatively unstable, capable of being dehydrated and of combining with other molecules. At higher temperatures they caramelize and char. It has been known for over a century that many sugars formed from formaldehyde in alkaline solution. Several alternative, synthetic pathways and reaction conditions have been proposed but specific details probably never be established. Condensation of formaldehyde, a prebiotic constituent, led to formation of glycol-aldehyde, trioses, tetroses, pentoses and hexoses in the laboratory. UV light, electric discharges and ionization radiation at the right pH and temperature may also promote synthesis of sugars. The phosphorylation of sugars is an indispensable step in the assembly of the building blocks of living matter. Cyanogen, a likely prebiotic constituent, is capable of catalyzing the synthesis of glucose mono- and diphosphate from glucose and orthophosphate.

Carbohydrates comprise a comprehensive group of naturally occurring substances, which include innumerable sugars and sugar derivatives, as well as high-molecular weight carbohydrates (polysaccharides) like starch and cellulose in plants and glycogen in animals. A polysaccharide molecule is composed of a large number of sugar or sugar-like units. Carbohydrates are of great importance in biology. The unique reaction, which makes life possible on Earth, namely the assimilation of the green plants, produces sugar, from which originate, not only all carbohydrates but, indirectly, also all other components of living organisms.

The important role of carbohydrates, generally, in the metabolism of living organisms, is well known. The biological breakdown of carbohydrates (often spoken of as "combustion") supplies the principal part of the energy that every organism needs for various vital processes. It is not surprising; therefore, that the carbohydrates and their metabolism have been the subject of comprehensive and in many respects successful biochemical and medical research for a long time.

This chapter is created in view of giving an idea about sugars (monosaccharides to polysaccharides; their derivatives, their important reactions, structures, function and biological importance. Efforts have made to explain the chemistry and organization of biomolecules in terms of stereioisomerism, optical isomerism, anomeric forms,

mutarotations etc. Information about polysaccharides in plants and living organisms, blood sugars and rare sugars is also included.

Definition

Carbohydrates are polyhydroxylated aldehydes or ketones and their derivatives. The word "carbohydrate" includes polymers and other compounds synthesized from polyhydroxylated aldehydes and ketones. They can be synthesized in the laboratory or in living cells. Simple carbohydrates or the entire carbohydrate family may also be called saccharides. In general carbohydrates have the empirical formula $(CH_2O)_n$ (Fig. 1). The term generated from carbon and hydrate; though some also contain nitrogen, phosphorus, or sulfur. Chemically, carbohydrates are molecules that are composed of carbon, along with hydrogen and oxygen - usually in the same ratio as that found in water (H_2O).

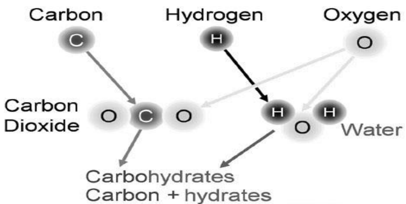


Fig. 1: Carbohydrate

They originate as products of photosynthesis, an endothermic reductive condensation of carbon dioxide requiring light energy and the pigment chlorophyll.

$$nCO_2 + nH_2O + \text{energy} \longrightarrow C_nH_{2n}O_n + nO_2$$

Typical carbohydrates are composed of strings or chains of monosaccharides - that is, chains of individual sugars.

Importance of carbohydrates

Carbohydrates are of great importance in biology. The unique reaction, which makes life possible on the Earth, namely the assimilation of the green plants, produces sugar, from which originate, not only all carbohydrates but, directly or indirectly, all other components of living organisms. The carbohydrates are a major source of metabolic energy, both for plants and for animals that depend on plants for food. Aside from the sugars and starch that meet this vital nutritional role, carbohydrates also serve as a structural material (cellulose), a component of the energy transport compound ATP, recognition sites on cell surfaces, and one of three essential components of DNA and RNA. Importance can be considered under following headings;

Metabolic/Nutritional

The important role of carbohydrates, generally, in the metabolism of living organisms, is well known. The biological breakdown of carbohydrates (often spoken of as "combustion") supplies the principal part of the energy that every organism needs for various processes.

Carbohydrates and their metabolism has been the subject of biochemical and medical research for a long time.

Carbohydrates play a major role in promoting health fitness, form a major part of food and help a great deal in building body strength, by generating energy. They are one among the three prominent macronutrients that serve as excellent energy providers, the other two being fats and proteins. Carbohydrate intake can take place in different forms like sugar, starch, fibers etc., which are a dietary staple in most parts of the world, and the oxidation of carbohydrates is the central energy-yielding pathway in most nonphotosynthetic organisms.

The functions of carbohydrates are multiple and it is owing to this fact that it becomes all the more necessary to incorporate carbohydrates in our meal. For instant for energy generation, sugars and starch act as the perfect fuel that enables us to carry out our physical activities efficiently and effectively. Fiber does wonders in keeping your bowel function going smooth. Carbohydrates add on to the taste and appearance of food item, thus making the dish tempting and mouthwatering. They are sometimes used as flavors and sweeteners. Carbohydrates aid in regulating blood glucose and also do good to our body by breaking down fatty acids, thus preventing ketosis. Talking about the importance of carbohydrates, apart from its direct benefits, there is also an added advantage of carbohydrate consumption and that is that carbohydrates are found in different foods, which if eaten, also pave way for consuming other essential nutrients. Therefore, it is preferable to go in for distinctive carbohydrate food sources.

Structural

The structural diversity possible by linking the different, common sugar is immense: theoretically far greater than that of proteins, which largely consist of 22 amino acids linked by a single type of peptide bond. Linkages between sugars can occur through a glycoside linkage between the anomeric, first carbon of a sugar in either α or β configuration with any of a variety of hydroxyl groups on the adjacent sugar. In fact, many possible combinations of sugars do not seem to exist.

Insoluble carbohydrate polymers serve as structural and protective elements in the cell walls of bacteria and plants and in the connective tissues of animals. Plant cell walls are complex arrangements of cellulose, hemicellulose and lignin. This contributes significantly to the overall digestibility of the fiber. The proportion of each component depends on species and age of the plant. Cellulose is the primary structural component of plants. It is found primarily in the cells walls and is a primary fiber component of animal feeds.

Communication

Glycosaminoglycans as polymers of derivatives of carbohydrates are of critical importance in intercellular communication in organisms. This ubiquitous class of linear polyanions interact with a wide variety of proteins, including growth factors and chemokines, which regulate important physiological processes. The presence of glycosaminoglycans on cell membranes and in the extracellular matrix also has resulted in their exploitation by infectious pathogens to gain access and entry into animal cells. Other carbohydrate polymers lubricate skeletal joints and participate in recognition and adhesion between cells.

More complex carbohydrate polymers covalently attached to proteins or lipids act as signals that determine the intracellular location or metabolic fate of these hybrid molecules, called glycoconjugates. Glycoconjugates carry various important functions of cell. Glycoproteins act as receptors and integral membrane proteins in membranes, cytoskeletal proteins in cytoplasm, extracellular proteins such as antibodies, hormones, collagen (found outside the cell), enzymes (RNase, DNase, lipases, cholinesterase, phosphatase, pepsinogen, glycosyltransferases) etc. Functions and locations of glycolipids within the cell are diverse—take for example: GAGs: submaxillary secretions, human gastric mucin, RBC membrane sialoglycoprotein, membrane protein, secretory proteins without enzymatic functions, immunoglobins, enzymes.

Biosynthesis of other compounds

Carbohydates are source of carbon for biosynthesis of other compounds. Most people know that the body uses carbohydrates for energy. For example, the simple carbohydrate glucose (dextrose) gets oxidized by liver cells. In exchange, the cells produce adenosine triphosphate (ATP), the main energy-providing compound in the cell. However, carbohydrates are used in a number of ways by plants, animals and bacteria, not just for energy.

In carbohydrate anabolism, simple organic acids can be converted into monosaccharides such as glucose and then used to assemble polysaccharides such as starch.

In living cells, glucose is oxidized to give carbon dioxide and energy. The energy produced is in the form of ATP and heat. This reaction is just like the combustion.

$$\text{Glucose} + O_2 \rightarrow CO_2 + \text{energy}$$

During photosynthesis in plants, glucose can be produced from CO2 and energy:

$$CO_2 + \text{energy} \rightarrow \text{glucose} + O_2$$

Here, the energy consumed is in the form of sunlight. This reaction is the reverse of the combustion reaction.

Classification

Carbohydrates are called saccharides or, if they are relatively small, sugars. Classifications of carbohydrates are outlined in the following table.

Table 1

Complexity	Simple Carbohydrates monosaccharides		Complex Carbohydrates disaccharides, oligosaccharides & polysaccharides			
Size	Tetrose C4 sugars	Pentose C5 sugars	Hexose C6 sugars	Heptose C7 sugars	Octose C8 sugars	Nonose C9 sugars
C=O Function	Aldose: sugars having an aldehyde function or an acetal equivalent.					
C=O Function	Ketose: sugars having a ketone function or an acetal equivalent.					

The compounds carbohydrates have common same functional groups, glyceraldehydes and gulose are classifed as aldoses and ribulose and dihydroxyacetone as ketoses (Fig. 2). All of these compounds are alcohols with many hydroxyl groups. They are polyhydroxylated and either aldehydes or ketones.

Fig. 2: Aldose & Ketose

A. Monosaccharides

The simplest and smallest unit of the carbohydrates is the monosaccharide, (mono = one, saccharide = sugar) from which disaccharides, oligosaccharides, and polysaccharides are constructed. Monosaccharides are either aldehydes or ketones, with one or more hydroxyl groups; the six-carbon monosaccharides glucose (an aldohexose) and fructose (a keto hexose) have five hydroxyl groups. (Fig. 3) The carbon atoms, to which hydroxyl groups are attached, are often chiral centers, and stereoisomerism is common among monosaccharides.

Fig. 3: Monosaccharide

Simple monosaccharides with four, five, six, and seven carbon atoms are called tetroses, pentoses, hexoses, and heptoses, respectively (Table1). Because these molecules have multiple asymmetric carbons, they exist as diastereoisomers, isomers that are not mirror images of each other, as well as enantiomers. In regard to these monosaccharides, the symbols D and L designate the absolute configuration of the asymmetric carbon farthest from the aldehyde or keto group. D-Ribose, the carbohydrate component of RNA, is a five-carbon aldose. D-Glucose, D-mannose, and D-galactose are abundant six-carbon aldoses. It may be noted that D-glucose and D-mannose differ in configuration only at C-2. Sugars differing in configuration at a single asymmetric center are called epimers. Thus, D-glucose and D-mannose are epimeric at C-2; D-glucose and D-galactose are epimers with respect to C-4 (Fig. 4)**.** *(More examples of monosaccharide will appear subsequently later)*

Fig. 4: Epimers

B. Disaccharides

A disaccharide consists of two monosaccharides joined by an O-glycosidic bond. Disaccharides can be homo- and heterodisaccharide (Fig. 5). Three most abundant disaccharides are sucrose, lactose, and maltose. In sucrose the anomeric carbon atoms of a glucose unit and a fructose unit are joined. Lactose, the disaccharide of milk, consists of galactose joined to glucose by a β (1→4) glycosidic linkage. In maltose, α (1→4) glycosidic linkage joins two glucose units. Sucrose and lactose are heterosaccharides and maltose is homosaccharide.

C. Oligosaccharide

An oligosaccharide is a saccharide polymer containing a small number (typically three to ten) of component sugars, and is also known as simple sugars. They are generally found either O- or N-linked to compatible amino acid side chains in proteins or to lipid moieties. They (homo-and hetero-oligosaccharides) are also liberated as intermediate products of saccharification by action of glycosidases on polysaccharides.

D. Polysaccharides

Polysaccharides are relatively complex carbohydrates. They are polymers made up of many monosaccharides joined together by glycosidic bonds. They are, therefore, very large, often branched, macromolecules. They tend to be amorphous, insoluble in water, and have no

sweet taste. When all the monosaccharides in a polysaccharide are of the same type, the polysaccharide is called a homopolysaccharide and when more than one type of monosaccharide is present, they are called heteropolysaccharides. Examples include storage polysaccharides such as starch and glycogen and structural polysaccharides such as cellulose, and chitin. Xylan a hemicellulose is a heteropolysaccharide. Polysaccharides have a general formula of $C_n(H_2O)_{n-1}$ where n can be any number between 200 and 2500. Considering that the repeating units in the polymer backbone are often six-carbon monosaccharides, the general formula can also be represented as $(C_6H_{10}O_5)_n$ where n={40...3000}.

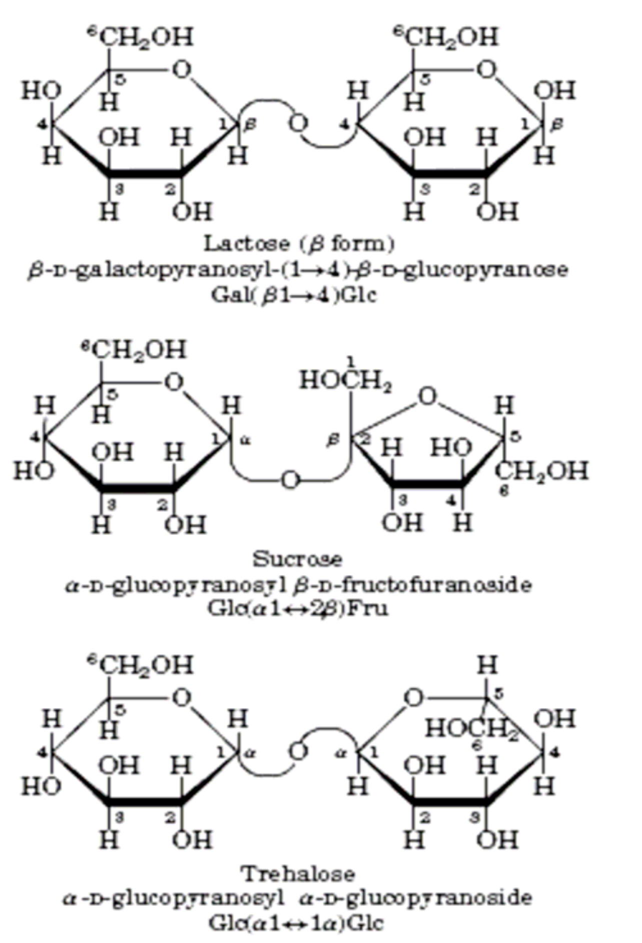


Fig. 5: Disaccharide

E. Families of Monosaccharides

Monosaccharides are colorless, crystalline solids that are freely soluble in water but insoluble in nonpolar solvents. Most have a sweet taste. The backbone of monosaccharides is an unbranched carbon chain in which all the carbon atoms are linked by single bonds. 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, the monosaccharide is an aldehyde and is called an aldose; if the carbonyl group is at any other position, the monosaccharide is a ketone and is called a ketose. The simplest monosaccharides are the two three-carbon trioses: glyceraldehyde, an aldose, and dihydroxyacetone, a ketose.

F. Structure of monosaccharide

Structures of monosaccharides are defined by Fischer projection, where all horizontal bonds project toward the viewer, while vertical bonds project away from the viewer. Therefore, a Fischer projection cannot be rotated by (2n+1)×90° in the plane of the page or the screen, as the orientation of bonds relative to one another can change, converting a molecule to its enantiomer (Fig. 4).

Examples of monosaccharide include glucose (dextrose), fructose, galactose, and ribose (Fig. 15). Monosaccharides are the building blocks of disaccharides like sucrose (common sugar) and polysaccharides (such as cellulose and starch and hemicellulose). Further, each carbon atom that supports a hydroxyl group (except for the first and last) is chiral, giving rise to a number of isomeric forms, all with the same chemical formula. Galactose and glucose are both aldohexoses (Fig. 4), but they have different chemical and physical properties.

With few exceptions (e.g., deoxyribose or dexyglucose), monosaccharides have the chemical formula $(CH_2O)_n$ + m with the chemical structure H $(CHOH)_nC{=}O(CHOH)_mH$. If n or m is zero, it is an aldehyde and is termed an aldose; otherwise it is a ketone and is termed a ketose. Examples of some aldo and keto sugars are presented in Table 2. Monosaccharides contain a ketone or aldehyde functional group, and hydroxyl groups on most or all of the non-carbonyl carbon atoms. Most monosaccharides form cyclic structures, which predominate in aqueous solution, by forming hemiacetals or hemiketals (depending on whether they are aldoses or ketoses) between an alcohol and the carbonyl group of the same sugar. Glucose, for example, readily forms a hemiacetal linkage between its carbon-1 and the hydroxyl group of its carbon-5. Since such a reaction introduces an additional stereogenic center, two anomers are formed (α-anomer and β-anomer) from each distinct straight-chain monosaccharide.

Table 2: Monosaccharides

Monosaccharides	Aldo form	Keto form
Trioses	Glyceraldehydes	Dihydroxyacetone
Tetroses	Erythrose and Threose	Erythrulose
Pentoses	Arabinose, Lyxose, Ribose and Xylose	Ribulose and Xylulose
Hexoses	Allose, Altrose, Galactose, Glucose, Gulose, Idose, Mannose and Talose	Fructose, Psicose, Sorbose and Tagatose
Heptoses	NA	Mannoheptulose, Sedoheptulose
Octoses	NA	2-Keto-3-deoxy-manno-octonate
Nonoses	NA	Sialose

The cyclic structure of monosaccharides is represented by **Haworth projection**, where the α-anomer has the OH- of the anomeric carbon under the ring structure, and the β-isomer, has the OH- of the anomeric carbon on top of the ring structure (Fig. 6).

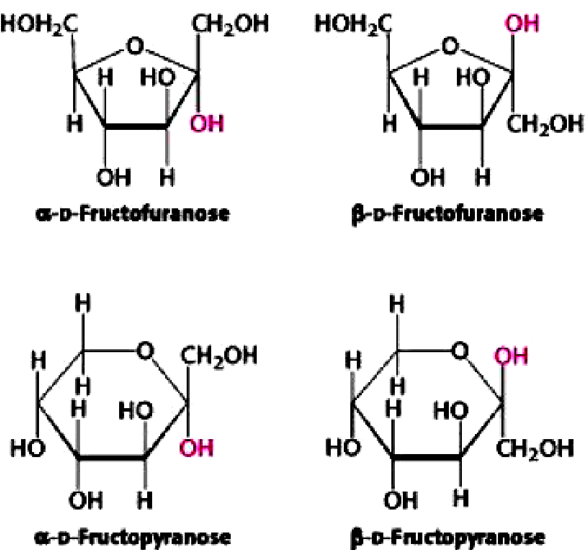


Fig. 6:Cyclic form

Haworth Projection: A Haworth projection is a common way of representing the cyclic structure of monosaccharides with a simple three-dimensional perspective. The Haworth projection was named after the English chemist Sir Walter N. Haworth.

A Haworth projection has the following characteristics:

- Carbon is the implicit type of atom. In the example on the right, the atoms numbered from 1 to 6 are all carbon atoms. Carbon 1 is known as the Anomeric Carbon.
- Hydrogen atoms on carbon are implicit. In the example, atoms 1 to 6 have extra hydrogen atoms not depicted.

A thicker line indicates atoms that are closer to the observer. In the following figure, atoms 2 and 3 (and their corresponding OH groups) are the closest to the observer, atoms 1 and 4 are further from the observer and finally the remaining atoms (5, etc.) are the furthest.

There is another way of representation of cyclic structure of saccharides; one is chair and another is boat conformation. In chair conformation, the α-anomer has the OH- of the anomeric carbon in an axial position, whereas the β-isomer has the OH- of the anomeric carbon in equatorial position The lowest-energy chair conformation, 6 of the 12 hydrogens are in axial positions their C-H bonds are parallel to each other and appear to stick up and down from the ring structure, the other 6 are in equatorial positions.

On the other hand, boat conformation has a higher energy than the chair form due to steric strain resulting from the two axial 1,4-hydrogen atoms. The torsional strain in the boat

conformation has a maximum value because all the carbon bonds are eclipsed. The boat and envelope forms are transition states between the twist forms and the twist and chair forms respectively, and are impossible to isolate. The twist-boat conformation is less stable than the chair conformation. The difference in energy between the chair and the twist-boat conformation of monosaccharides can be measured indirectly by taking the difference in activation energy for the conversion of the chair to the twist-boat conformation and that of the reverse isomerization (Fig. 7).

Chair form Boat form

Fig. 7: Chair and Boat form

Axial & equatorial position: The axial hydrogens are those sigma bonds that are parallel to an imaginary axis running through the ring structure (Scheme 1a). The equatorial hydrogens are those whose sigma bonds are perpendicular to that axis (Scheme 1b).

Scheme 1a Scheme 1b

In the figures given not all of the hydrogens attached to the ring carbons in cyclohexane are equivalent. Replacing all but one of the 12 hydrogens with deuterium leaving one position as regular hydrogen can prove this. Because the ring flip occurs so rapidly at room temperature, there will be only one nuclear magnetic resonance (NMR) signal for the one hydrogen. Deuterium atoms do not have a spin state that would generate an NMR signal. However when we slow the molecular motion down by lowering the temperature of the sample to an extremely low temperature so the ring flip occurs much less rapidly, we observe two separate NMR signals with slightly different chemical shifts. This would indicate that the hydrogen has changed positions as the ring underwent flipping so that the two positions are non-equivalent. These two non-equivalent hydrogen positions are called the axial and equatorial positions. The axial bonds run vertical to the molecule. The equatorial bonds run approximately horizontal to the molecule. During the ring flip process the axial hydrogens become equatorial and the equatorial hydrogens become axial. (Scheme 1b above)

Isomerism

Isomers are defined as molecule with the same chemical formula and often with the same kinds of chemical bonds between atoms, but in which the atoms are arranged differently (analogous to a chemical anagram). That is to say, they have different structural formulae. Many isomers share similar if not identical properties in most chemical contexts. Isomerism can be of two types: Structural isomerism and Stereoisomerism

A. Chirality

A carbon is chiral if it is sp^3 hybridized and it has 4 different groups attached to it. Molecule is chiral if it is not superimposable on its mirror image. This means that none of its conformations (out of the infinite number of them) has a plane of symmetry through it. Remember, there are as many conformations in a molecule as there are rotations about single bonds.

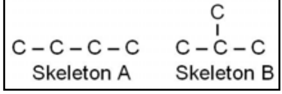
Chiral carbon centers

All chiral molecules have nonsuperimposable mirror images. And as a general rule of thumb, chiral molecules must have one or more chiral centers -- that is, carbons that have four non-identical substituents around it. (There are, of course, exceptions to this rule). A classic case of a simple chiral molecule is the following halogenated methane derivative to it. Chiral carbons result in stereoisomers.

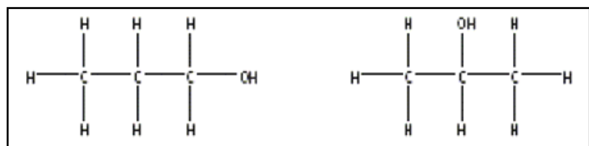
B. Structural isomerism

In structural isomers, the atoms and functional groups are joined together in different ways. This group includes chain isomerism, position isomerism and functional group isomerism. This is also termed as constitutional isomerism.

Chain isomerism – Here hydrocarbon chains have variable amounts of branching. In carbon-chain isomerism, the difference between the isomers is the length of the carbon chain. For example, if the compound contains four carbon atoms, they can be arranged with two different chain lengths as below:



Positional isomerism – This deals with the position of a functional group on a chain. Positional isomers have the same functional group in different locations on the carbon skeleton. For example, there are two isomers of molecular formula C_3H_8O. In one isomer (propan-1-ol) the –OH group is bonded to an end carbon atom; in the other (propan-2-ol) it is bonded to the middle carbon atom.

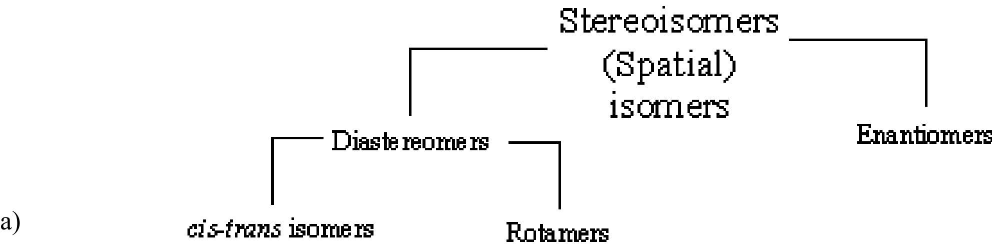


Functional-group isomerism – Here one functional group is split up into different ones. The isomers are members of different homologous series and, therefore, have different functional groups. For example, there are two isomers of molecular formula C_2H_6O. One isomer is the alcohol, ethanol, CH_3CH_2OH; the other is the ether, methoxymethane, CH_3OCH_3.

$CH_3 - O - CH_3$	$CH_3 - CH_2 - OH$
Di-Methyl Ether	Ethanol

C. Stereoisomerism

The stereoisomers have same chemical formula and connectivity of atoms, but different arrangements of those atoms. This is also termed as spatial isomerism. This class includes enantiomers and diastereomers. Diastereomerism is of two types: conformational isomerism (conformers) and *cis-trans* isomerism . Conformers can be referred to as having a diastereomeric relationship; the isomers over all are not diastereomers, since bonds in conformers can be rotated to make them mirror images.



Enantiomers: Enantiomers are mirror images of chiral molecules. Enantiomers are not superimposable on each other. Only chiral molecules will have a nonsuperimposable mirror image. If two molecules are enantiomers, all of the chiral centers in it must have opposite configurations.

The above carbon atom has four non-identical substituents around it, making this carbon a chiral center, and as proof of its chirality the molecule has a non-superimposable mirror image. A fancy term used in textbooks and in the literature to describe molecules that are mirror images of each other are enantiomers, as in "the enantiomer of the left molecule above is the molecule on the right, its mirror image."
To distinguish between enantiomers, chemists use the R and S classification system. Stereocenters, (sometimes called chiral centers or stereogenic centers) are carbons that have four non-identical substituents on them, and are designated as either of R stereochemistry or S stereochemistry. If a molecule has one stereocenter of R configuration, then in the mirror image of that molecule, the stereocenter would be of S configuration, and *vice versa* (Fig. 8).

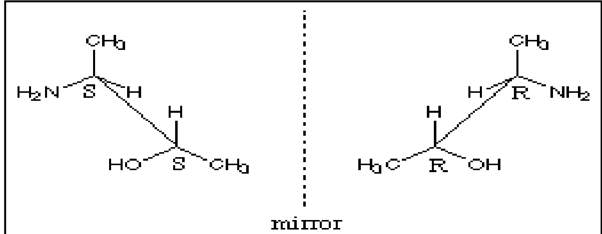


Fig. 8: Enantiomers

Enantiomers have identical physical properties (b.p., m.p., density, etc.) with several exceptions such as enantiomers have opposite specific rotations; they are distiguishable by NMR spectroscopy in chiral media (but not achiral), they may exhibit different chemical properties, such as very different rates of reaction.

Given the importance of stereochemistry in reactions between biomolecules, biochemists must name and represent the structure of each biomolecule so that its stereochemistry is unambiguous. For compounds with more than one chiral center, the most useful system of nomenclature is the RS system. In this system, each group attached to a chiral carbon is assigned a priority. The priorities of some common substituents are

OOCH2 > OOH> ONH2 > OCOOH > OCHO> OCH2OH > OCH3 > OH

For naming in the RS system, the chiral atom is viewed with the group of lowest priority (4 in the diagram on the next page) pointing away from the viewer. If the priority of the other three groups (1 to 3) decreases in clockwise order, the configuration is (R) (Latin *rectus*, "right"); if in counterclockwise order, the configuration is (S) (Latin *sinister*, "left"). In this way each chiral carbon is designated either (R) or (S), and the inclusion of these designations in the name of the compound provides an unambiguous description of the stereochemistry at each chiral center.

Determination of R or S configuration can be applied in three steps:

1.Ordering of the substituents coming off the stereogenic carbon atom using the Cahn-Ingold-Prelog rules.

[The Cahn-Ingold-Prelog priority rules, CIP system or CIP conventions are a set of rules used in organic chemistry to name the stereoisomers of a molecule. A molecule may contain any number of stereocenters and any number of double bonds, and each give rise to two possible configurations. The purpose of the CIP system is to assign an R or S descriptor to each stereocenter and an E or Z descriptor to each double bond so that the configuration of the entire molecule can be specified uniquely by including the descriptors in its systematic name.]

2. Rotation of the molecule until the lowest priority substituent is in the back

3. Depiction of a curve through the substituents in descending order of priority. If the curve is clockwise, the stereocenter is of R configuration. If the curve is counterclockwise, the stereocenter is of S configuration.

1. Ordering the substituents

The substituents with higher molecular weight atoms are given a higher priority number. In our example below, iodine would be 1, bromine 2, chlorine 3, and fluorine 4, because iodine has the highest molecular weight (and is therefore highest priority) and fluorine has the lowest molecular weight (and is therefore the lowest priority).

If the first atom of two substituents happen to be identical in molecular weight, go to the next atom and make the molecular weight comparison (eg. an ethyl group would have higher priority over a methyl group). Tthis applies in the same fashion to carbonyls, C=O, and imines, C=N. A carbon with a double bond to another carbon is treated as a carbon singly bonded to two carbons, as shown below. This means that, for example, an ethylene substituent, R-CH=CH2, will have a higher priority then an ethyl substituent (R-CH2CH3). Shown here is an ethylene substituent (often called an allyl substituent). By the Cahn-Ingold-Prelog rules for assigning R and S nomenclature, this allyl group can be redrawn with each double bond carbon singly bonded to an additional carbon with three "phantom ligands," that are ignored.

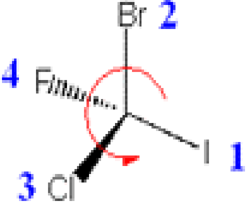
Our example molecule can now be numbered as follows:

2. Rotation of the molecule

There are a couple of different ways to go from the priority numbering to determining R and S configuration. One of the best methods is to rotate the molecule until the number 4 priority substituent is in the back, as shown below.

3. Draw the curve

A curve is then drawn from the 1 to 2 to 3 priority substituents, ignoring the 4th priority subsituent (as shown below). If that curve goes clockwise then that stereocenter is of the R configuration. If the curve goes in a counterclockwise direction, then that stereocenter is of S configuration. In our example below, the curve goes counterclockwise so the stereocenter is of S configuration.



The name of the compound above, then, would be (S)-bromo, chloro, fluoro, iodomethane, and the name of its enantiomer, or its mirror image, would be (R)-bromo, chloro, fluoro, iodomethane.

b) **Diastereomers**: Diastereomers are stereoisomers that are not enantiomers. If two molecules differ only in the configuration of some of their stereocenters but all stereocenters are not the opposite configuration, then the molecules are diastereomers (Fig. 9). In this class molecules do not have to be chiral to have a diastereomer, but they do have to contain at least two chiral centers.

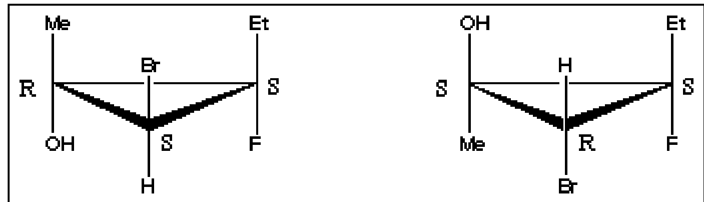


Fig. 9: Diasteromers

i) **Conformational diastereomers:** Conformational isomers are different conformers of a diastereomer molecule. They differ only in how groups are rotated about single bonds. Because at room temperature there is free rotation about single bonds, the conformation of a given molecule will be changing constantly. There are an infinite number of conformations for a given molecule as a result of an infinite number of posible rotations. A molecule does not have to be chiral or contain chiral centers to have conformational isomers (Fig. 10).

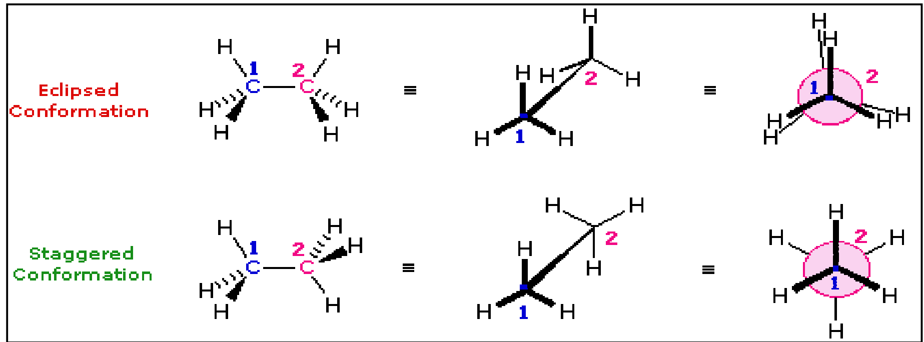


Fig. 10: Conformational diasteromers

ii) Cis-Trans diastereomers: Cis-trans isomerism or geometric isomerism form of stereoisomerism relates to the orientation of functional groups within a molecule. Such isomers contain double bonds, which cannot rotate, but they can also arise from ring structures, wherein the rotation of bonds is greatly restricted (Fig. 11).

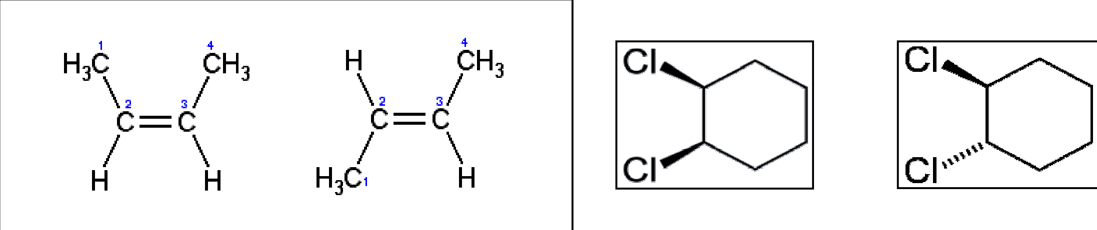


Fig. 11: *Cis-Trans* diasteromers

There are two forms of a *cis-trans* isomer, the *cis* and *trans* versions. When the substituent groups are oriented in the same direction, the diastereomer is referred to as *cis*, whereas, when the substituents are oriented in opposing directions, the diastereomer is referred to as *trans*. An example of a small hydrocarbon displaying cis-trans isomerism is 2-butene. Alicyclic compounds can also display cis-trans isomerism. As an example of a geometric isomer due to a ring structure, consider 1,2-dichlorocyclohexane.

Structural isomers have different chemical properties; but stereoisomers behave identically in most chemical reactions. Enzymes differentiate between stereoisomers of a compound, and organisms do prefer one stereoisomer to the other. Some stereoisomers also differ in the way they rotate polarized light.

There are other classes of isomers as below:

Tautomers are isomers of the same chemical substance that spontaneously convert to each other, even when pure. They have different chemical properties, but cannot be isolated from each other.

D. Mutarotation

The term for interconversion between the two anomers is called mutarotation. The two different anomers are two distinct chemical structures, and thus have different physical and chemical properties, notably optical rotation. For example, α-D-glucose has an optical rotation of +112 degrees and its anomer, β-D-glucose, has an optical rotation of +19 degrees.

Isomeric forms of monosaccharides that differ only in their configuration about the hemiacetal or hemiketal carbon atom are called anomers. The hemiacetal (or carbonyl) carbon atom is called the anomeric carbon. The α and β anomers of D-glucose interconvert in aqueous solution by a process called mutarotation (Fig. 12).

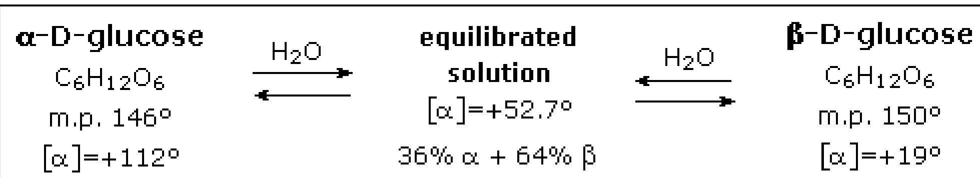


Fig. 12: Mutarotation

The figure is self-explanatory and shows that 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 (Fig. 13).

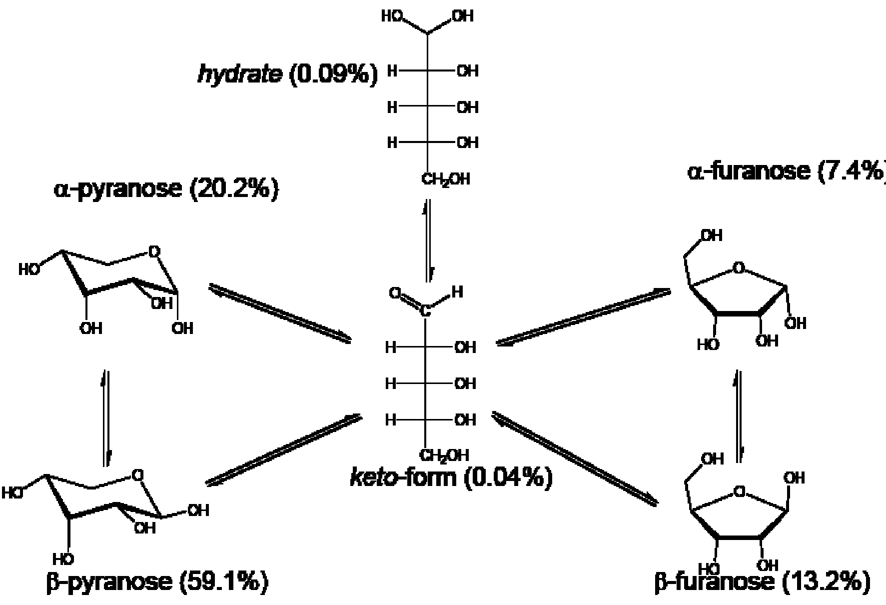


Fig. 13: Mutarotate forms of glucose

Anomers

In sugar chemistry, an anomer is a special type of epimer. It is a stereoisomer (diastereomer, more exactly) of a saccharide (in the cyclic form) that differs only in its configuration at the hemiacetal (or hemiketal) carbon, also called the anomeric carbon. If the structure is analogous to one with the hydroxyl group on the anomeric carbon in the axial position of glucose, then the sugar is an alpha anomer. If, however, that hydroxyl is in the equatorial position, then the sugar is a beta anomer. For example, α-D-glucopyranose and β-D-glucopyranose, the two cyclic forms of glucose, are anomers. For -L- the α and β are contrarywise. The anomeric effect helps stabilize the α-anomer.

α-D-glucopyranose ⇌ β-D-glucopyranose

[The α- and β-anomers of D- glucopyranose]

E. Configuration of glucose

There are four chiral centers in glucose, which indicate possibility of sixteen stereoisomers that would exist as eight diastereomeric pairs of enantiomers. German chemists Emil Fischer in 1891 devised a method of representing the configuration of each chiral center in an unambiguous manner. Fischer established a network of related aldose configurations termed as D-family through an arbitrary choice for (+)-glucose. Consequently, The mirror images of these configurations were designated the L-family of aldoses.

The last chiral center in an aldose chain (farthest from the aldehyde group) was chosen by Fischer as the D / L designator site. If the hydroxyl group in the projection formula pointed

to the right, it was defined as a member of the D-family and a left directed hydroxyl group (the mirror image) then represented the L-family. The sign of a compound's specific rotation (an experimental number) does not correlate with its configuration (D or L) and optical rotation can be measured by with a polarimeter.

Three-dimensional structure is described by Configuration and Conformation

The covalent bonds and functional groups of a biomolecule are, of course, central to its function, but so also is the arrangement of the molecule's constituent atoms in three-dimensional space—its stereochemistry. A carbon-containing compound commonly exists as stereoisomers, molecules with the same chemical bonds but different stereochemistry—that is, different configuration, the fixed spatial arrangement of atoms. Interactions between biomolecules are invariably stereospecific, requiring specific stereochemistry in the interacting molecules.

Configuration is conferred by the presence of either (1) double bonds, around which there is no freedom of rotation, or (2) chiral centers, around which substituent groups are arranged in a specific sequence. The dentifying characteristic of configurational isomers is that they cannot be interconverted without temporarily breaking one or more covalent bonds. A binding site (on an enzyme, for example) that is complementary to one of these molecules would not be a suitable binding site for the other, which explains why the two compounds have distinct biological roles despite their similar chemistry.

In the second type of configurational isomer, four different substituents bonded to a tetrahedral carbon atom may be arranged two different ways in space—that is, have two configurations —yielding two stereoisomers with similar or identical chemical properties but differing in certain physical and biological properties. A carbon atom with four different substituents is said to be asymmetric, and asymmetric carbons are called chiral centers (Greek chiros, "hand"; some stereoisomers are related structurally as the right hand is to the left). A molecule with only one chiral carbon can have two stereoisomers; when two or more (n) chiral carbons are present, there can be 2n stereoisomers. Some stereoisomers are enantiomers and others, pairs of stereoisomers that are not mirror images of each other are called diastereomer.

Distinct from configuration is molecular conformation, the spatial arrangement of substituent groups that, without breaking any bonds, are free to assume different positions in space because of the freedom of rotation about single bonds. In the simple hydrocarbon ethane, for example, there is nearly complete freedom of rotation around the COC bond. Many different, interconvertible conformations of ethane are possible, depending on the degree of rotation. Two conformations are of special interest: the staggered, which is more stable than all others and thus predominates, and the eclipsed, which are least stable. We cannot isolate either of these conformational forms, because they are freely interconvertible. However, when one or more of the hydrogen atoms on each carbon is replaced by a functional group that is either very large or electrically charged, freedom of rotation around the COC bond is hindered. This limits the number of stable conformations of the ethane derivative.

Interactions between biomolecules are stereospecific

Biological interactions between molecules are stereospecific: the “fit” in such interactions must be stereochemically correct. The three-dimensional structure of biomolecules large and small—the combination of configuration and conformation—is of the utmost importance in their biological interactions: reactant with enzyme, hormone with its receptor on a cell surface, antigen with its specific antibody. The study of biomolecular stereochemistry with precise physical methods is an important part of modern research on cell structure and biochemical function. In living organisms, chiral molecules are usually present in only one of their chiral forms. For example, the amino acids in proteins occur only as their L isomers; glucose occurs only as its D isomer. In contrast, when a compound with an asymmetric carbon atom is chemically synthesized in the laboratory, the reaction usually produces all possible chiral forms: a mixture of the D and L forms, for example. Living cells produce only one chiral form of biomolecules because the enzymes that synthesize them are also chiral.

Stereospecificity, the ability to distinguish between stereoisomers, is a property of enzymes and other proteins and a characteristic feature of the molecular logic of living cells. If the binding site on a protein is complementary to one isomer of a chiral compound, it will not be complementary to the other isomer, for the same reason that a left glove does not fit a right hand.

Optical activity

Optical activity is the ability of a chiral molecule to rotate the plane of plane-polairsed light. It is measured using a polarimeter, which consists of a light source, polarising lens, sample tube and analysing lens. When light passes through a sample that can rotate plane polarised light, the light appears to dim to the eye because it no longer passes straight through the polarising filters. The amount of rotation is quantified as the number of degrees that the analysing lens must be rotated by so that it appears as if no dimming, of the light has occurred.

Louis Pasteur encountered the phenomenon of optical activity in 1843, during his investigation of the crystalline sediment that accumulated in wine casks (a form of tartaric acid called paratartaric acid—also called racemic acid, from Latin racemus, “bunch of grapes”). He used fine forceps to separate two types of crystals identical in shape but mirror images of each other. Both types proved to have all the chemical properties of tartaric acid, but in solution one type rotated polarized light to the left (levorotatory), the other to the right (dextro rotatory). Pasteur later described the experiment and its interpretation:

In isomeric bodies, the elements and the proportions in which they are combined are the same, only the arrangement of the atoms is different. We know, on the one hand, that the molecular arrangements of the two tartaric acids are asymmetric, and, on the other hand, that these arrangements are absolutely identical, excepting that they exhibit asymmetry in opposite directions. Are the atoms of the dextro acid grouped in the form of a right-handed spiral, or are they placed at the apex of an irregular tetrahedron, or are they disposed according to this or that asymmetric arrangement?

X-ray crystallographic studies in 1951 confirmed that the levorotatory and dextrorotatory forms of tartaric acid are mirror images of each other at the molecular level and established the absolute configuration of each.

Measurement of optical activity:
When rotation is quantified using a polarimeter it is known as an observed rotation, because rotation is affected by path length (l), or the distance for which the light travels through a sample and concentration (c), of the sample present to rotate the light. When these effects are eliminated a standard for comparison of all molecules is obtained, the specific rotation, [α], when concentration is expressed as sample in grams /100ml solution.

$$[\alpha] = 100a / cl$$

Specific rotation is a physical property like the boiling point of a sample. Enantiomers will rotate the plane of polarisation in exactly equal amounts (same magnitude) but in opposite directions. Dextrorotary designated as (+), clockwise rotation (to the right) Levorotary designated as (-), anti-clockwise rotation (to the left)

If only one enantiomer is present a sample is considered to be optically pure. When a sample consists of a mixture of enantiomers, the effect of each enantiomer cancels out, molecule for molecule. For example, a 50:50 mixture of two enantiomers or a racemic mixture will not rotate plane polarised light and is optically inactive. A mixture that contains one enantiomer excess, however, will display a net plane of polarisation in the direction characteristic of the enantiomer that is in excess.

Determinination of optical purity
The optical purity or the enantiomeric excess (ee %) of a sample can be determined as follows:

Optical purity = % enantiomeric excess = % enantiomer1 - % enantiomer2

$$= 100\ [\alpha]_{mixture} / [\alpha]_{pure\ sample}$$
$$ee\ \% = 100\ ([R]-[S]) / ([R]+[S])$$

where [R] = concentration of the R-isomer
[S] = concentration of the S-isomer

Important reactions of Aldehydes and Ketones

Carbohydrates being polyhydroxylated aldehydes and ketones, it is important to know about reactions of aldehyde and ketones, to understand reactivity of sugars.

A. Reversible Addition Reactions

a. Hydration and Hemiacetal Formation
Hydration of aldehydes leads to hemiacetal formation as shown below. It is well known that water adds rapidly to the carbonyl function of aldehydes and ketones. The resulting hydrate (hemiacetal) is unstable relative to the reactants and cannot be isolated. The addition products are called hemiacetals and are unstable.

$$R_2C{=}O \quad + R'OH \rightleftharpoons R'O{-}(R_2)\ C{-}O{-}H \quad \text{(a hemiacetal)}$$

b. Acetal Formation
Acetals are geminal-diether derivatives of aldehydes or ketones, formed by reaction with two equivalents of an alcohol and elimination of water. Ketone derivatives of this kind were

once called ketals. The following equation shows the overall stoichiometric change in acetal formation.

$$R_2C{=}O \;+ 2\,R'OH \rightarrow R_2C\,(OR')_2 \;+\; H_2O \quad \text{(an acetal)}$$

For an effective acetal formation two additional features, an acid catalyst must be used; and second, the water produced with the acetal must be removed from the reaction. The latter is important, since acetal formation is reversible. The acetals may be hydrolyzed back by treatment with aqueous acid. The mechanism shown here applies to both acetal formation and acetal hydrolysis by the principle of microscopic reversiblity.

Reaction: Mechanism of acetal formation

The importance of acetals as carbonyl derivatives lies chiefly in their stability and lack of reactivity in neutral to strongly basic environments. As long they do not get exposed to acids, especially aqueous acid, acetals show lack of reactivity similar to ethers in general.

Among the most useful and characteristic reactions of aldehydes and ketones is their reactivity toward strongly nucleophilic (and basic) metallo-hydride, alkyl and aryl reagents. If the carbonyl functional group is converted to an acetal these powerful reagents have no effect; thus, acetals are excellent protective groups, when these irreversible addition reactions must be prevented.

c. Formation of Imines and Related Compounds

Imine derivatives are formed by the reaction of aldehydes and ketones with ammonia or 1º-amines, also known as Schiff bases, (compounds having a C=N function). This reaction plays an important role in the synthesis of 2º-amines. Water is eliminated in the reaction, which is acid-catalyzed and reversible in the same sense as acetal formation.

$$R_2C{=}O \;+\; R'NH_2 \rightleftharpoons R'NH{-}(R_2)C{-}O{-}H \rightleftharpoons R_2C{=}NR' \;+\; H_2O$$

B. Irreversible Addition Reactions

Most hydrates and hemiacetals (Y = OH & OR), are known to decompose spontaneously to the corresponding carbonyl compounds. Aminols (Y = NHR) are intermediates in imine formation, and also revert to their carbonyl precursors if dehydration conditions are not employed. If substituent Y is an alkyl group or an aryl group, the resulting alcohol is a stable compound and does not decompose with loss of hydrogen or hydrocarbons, even on

heating. It follows then, that if nucleophilic reagents corresponding to H:(–), R:(–) or Ar:(–) add to aldehydes and ketones, the alcohol products of such additions will form irreversibly. Free anions of this kind would be extremely strong bases and nucleophiles, but their extraordinary reactivity would make them difficult to prepare and use.

a. Reduction by Complex Metal Hydrides

Addition of a hydride anion to an aldehyde or ketone would produce an alkoxide anion, which on protonation should yield the corresponding alcohol. Aldehydes would give 1º-alcohols and ketones would give 2º-alcohols.

RCH=O + H:(–) ⟶ RCH2O(–) + H3O(–) ⟶ RCH2OH

Reduction of α, β-unsaturated ketones by metal hydride reagents sometimes leads to a saturated alcohol, especially with sodium borohydride. This product is formed by an initial conjugate addition of hydride to the β-carbon atom, followed by ketonization of the *enol* product and reduction of the resulting saturated ketone (equation 1 below).

1) RCH=CHCOR'	+	NaBH4 (aq. alcohol)	——>	RCH=CHCH(OH)R'	+ RCH2-CH2CH(OH)R'
				1,2-addition product	1,4-addition product
2) RCH=CHCOR'	+	NaBH4 & CeCl3 -15º	——>	RCH=CHCH(OH)R'	
				1,2-addition product	

Reaction: Reduction of α, β-unsaturated ketones by metal hydride reagents

C. Other Carbonyl group reactions

a. Oxidation

The most common and characteristic oxidation reaction is the conversion of aldehydes to carboxylic acids. The carbon atom of a carbonyl group has a relatively high oxidation state. This is reflected in the fact that most of the reactions described thus far either cause no change in the oxidation state (e.g. acetal and imine formation) or effect a reduction (e.g. organometallic additions and deoxygenations). In the shorthand equation shown here the [O] symbol refers to unspecified oxidation conditions, which effect the desired change. Several different methods of accomplishing this transformation will be described here.

RCH=O + [O] ⟶ RC(OH)=O

In the oxidations of 1º and 2º-alcohols, aldehydes were converted to carboxylic acids, presumably via the hydrate. Useful tests for aldehydes, Tollens' test, Benedict's test & Fehling's test, take advantage of this ease of oxidation by using Ag(+) and Cu(2+) as oxidizing agents (oxidants).

RCH=O + 2 [Ag(+) OH(–)] ⟶ RC(OH)=O + 2 Ag (metallic mirror) + H2O

When silver cation is the oxidant, as in the above equation, it is reduced to metallic silver in the course of the reaction, and this deposits as a mirror on the inner surface of the reaction vessel. The Fehling and Benedict tests use cupric cation as the oxidant. This deep blue reagent is reduced to cuprous oxide, which precipitates as a red to yellow solid. All these cation oxidations are conducted under alkaline conditions. To avoid precipitation of the

insoluble metal hydroxides, the cations are stabilized as complexed ions. Silver is used as its ammonia complex, $Ag(NH_3)^{2(+)}$, and cupric ions are used as citrate or tartrate complexes.

Important reactions of sugars

A. Oxidation

Sugars are classified as reducing or non-reducing based on their reactivity with Tollens', Benedict's or Fehling's reagents. If a sugar is oxidized by these reagents, it is called reducing, since the oxidant (Ag^+ or Cu^{+2}) is reduced in the reaction, as evidenced by formation of a silver mirror or precipitation of cuprous oxide. The Tollens' test is commonly used to detect aldehyde functions; and because of the facile interconversion of ketoses and aldoses under the basic conditions of this test, ketoses such as fructose also react and are classified as reducing sugars.

When the aldehyde function of an aldose is oxidized to a carboxylic acid the product is called an aldonic acid. Because of the 2º hydroxyl functions that are also present in these compounds, a mild oxidizing agent such as hypobromite must be used for this conversion.

D-(-)-ribose —HOBr, pH ~ 8→ ribonic acid

D-(-)-ribose —HNO3 dilute→ ribaric acid (achiral) (mirror)

D-(-)-arabinose —HNO3 dilute→ D-arabinaric acid ≡ D-lyxaric acid ←HNO3 dilute— D-(-)-lyxose

Oxidation of non anomeric hydroxyl groups can be achieved by the following oxydising agents

a. Periodate oxidation

This reaction can be used quantitatively to measure the number of adjacent hydroxyl group in a molecule; this is used in determination of polysaccharide structure.

$2IO_4^-$ → + HCOOH

Aldose Alderic acid

b. Di nitrogen tetra oxide:
This reagent is specific for oxidation of primary alcohol groups. If we start with an aldose, we get an alduronic acid. In case of a disaccharide, an aldobiouronic acid is obtained.

N_2O_4

Methyl a-D- glucopyranoside Methyl a-D- glucopyranosiduronic acid

c. Hydrogen peroxide
It is a non-specific oxidant. It depolymerizes oligo- or polysaccharides and involves a free radical mechanism. It also employs an Fe^{+2} catalyst.

$$Fe^{+2} + HO\text{-}OH \longrightarrow Fe^{+3} + HO^{\circ} + OH^{-}$$

B. Reduction

Sodium borohydride reduction of an aldose makes the ends of the resulting alditol chain identical, $HOCH_2(CHOH)_nCH_2OH$, thereby accomplishing the same configurational change produced by oxidation to an aldaric acid. Thus, allitol and galactitol from reduction of allose and galactose are achiral, and altrose and talose are reduced to the same chiral alditol. A summary of these redox reactions and derivative nomenclature is given in the following.

Derivatives of $HOCH_2(CHOH)_nCHO$		
HOBr Oxidation	——>	$HO\text{-}CH_2(CHOH)_n\text{-}CO_2H$ an Aldonic Acid
HNO_3 Oxidation	——>	$H_2O\text{-}C(CHOH)_n\text{-}CO_2H$ an Aldaric Acid
$NaBH_4$ Reduction	——>	$HO\text{-}CH_2(CHOH)_n\text{-}CH_2OH$ an Alditol

Commercial examples

reduction

$NaBH_4$,
H_2/Pt or Pd,
H_2/Raney Ni

D-Glucose D-Glucitol (Sorbitol)

Sorbitol was discovered in the barriers of mountain ash (*Sorbus aucuparia*).
[**Importance**: Functions as humectant (water retainer), Non cariogenic (S. mutans can't metabolize). It is about 50% as sweet as sucrose, used in toothpaste, in sugarless gums, mints, hard candies, cough drops. Used as a cryoprotectant in surimi, as a starting material for sorbitan esters (emulsifiers) as well as for vitamin C synthesis].

reduction

D-fructose D-glucitol D-mannitol

[**Importance**: In the laboratory mannitol prepared by hydrogenation of D mannose, by fermentation; and commercially by hydrogenolysis of sucrose. It is not a humectant, crystallizes easily, moderately water soluble and 65% sweet as sucrose. It is used in making candies, sugar free chocolate, mints and cough drops. It is also used as an anticaking agent as well as dusting agent.]

D-xylose — hydrogenation → Xylitol

[**Importance**: It has an endothermic heat of solution, and, therefore, cooling sensation in mouth. It is 70% to 95% as sweet as sucrose and is non cariogenic.]

Maltose (from corn syrup) — hydrogenation → Maltitol

[**Importance**: It is used in making low calorie candies and miscellaneous other food items.]

C. Osazone Formation

The osazone reaction was developed and used by Emil Fischer to identify aldose sugars differing in configuration only at the alpha-carbon. The equation shows the general form of the osazone reaction, which effects an alpha-carbon oxidation with formation of a bis-phenylhydrazone, known as an osazone. Application of the osazone reaction to D-glucose and D-mannose demonstrates that these compounds differ in configuration only at C-2.

D-(+)-glucose — $3\ C_6H_5NHNH_2$ → common D-osazone ← $3\ C_6H_5NHNH_2$ — D-(+)-mannose

Reaction: Osazone Formation

D. Chain Shortening and Lengthening

These two procedures permit an aldose of a given size to be related to homologous smaller and larger aldoses. Ruff degradation of the pentose arabinose gives the tetrose erythrose. Working in the opposite direction, a Kiliani-Fischer synthesis applied to arabinose gives a mixture of glucose and mannose.

Using these reactions we can now follow Fischer's train of logic in assigning the configuration of D-glucose.

1. Ribose and arabinose (two well known pentoses) both gave erythrose on Ruff degradation. As expected, Kiliani-Fischer synthesis applied to erythrose gave a mixture of ribose and arabinose.
2. Oxidation of erythrose gave an achiral (optically inactive) aldaric acid. This defines the configuration of erythrose.
3. Oxidation of ribose gave an achiral (optically inactive) aldaric acid. This defines the configuration of both ribose and arabinose.
4. Ruff shortening of glucose gave arabinose, and Kiliani-Fischer synthesis applied to arabinose gave a mixture of glucose and mannose.
5. Glucose and mannose are, therefore, epimers at C-2, a fact confirmed by the common product from their osazone reactions.
6. A pair of structures for these epimers can be written, but which is glucose and which is mannose?

Note: Fischer looked for and discovered a second aldohexose that represented the end group exchange for the epimer lacking the latent C_2 symmetry. This compound was L-(+)-gulose, and its exchange relationship to D-(+)-glucose was demonstrated be oxidation to a common aldaric acid product. The remaining epimer is, therefore, mannose.

Reactions of sugar due to hydroxyl groups

Chemical properties of monosaccharides are a consequence of different levels of reactivity of primary and secondary alcohol groups. The alcohol group is nucleophilic, acidic (pKa 10-18), and is also easily oxidized by a wide range of reagents such as pyridinium dichromate (PDC) and chromium trioxide (CrO_3). A hydroxyl group can participate in numerous transformations under relatively mild conditions.

The reactivity of sugar molecules (free or functionalized) is heavily governed by the geometry of the molecule. The specific conformation of a sugar molecule (aldoses or ketoses) causes different reactivities of its secondary alcohol group.

A. Ester formation

Reaction type: Nucleophilic Acyl Substitution

Reaction: Acylation of -OH (preparation of esters)

Acetic anhydride in pyridine at 0ºC is used as a reagent for the above reaction. Esters are also formed using acetyl chloride, benzoyl chloride, diphenylphosphorochloridate, and p-toluenesulphonyl chloride.

Uses: Intermediates in the synthesis of glycosides and carbohydrate derivatives. Esters may be used in determination of the monosaccharide composition of polysaccharides.

B. Ether formation

Reaction type: Nucleophilic Substitution (SN2)

Reaction: Alkylation of -OH group (preparation of ethers)

Methylations occur smoothly with methyl iodide in dimethyl sulfoxide (DMSO) using DMSO anion as an acid acceptor. Reagents methyl iodide or dimethyl sulphate and an agent to promote ionization of the hydroxyl groups and to accept the acid released are used. The methyl ethers are stable under acidic and alkaline conditions. The -OR group at the acetal center can be converted back to an -OH with aqueous acid.

Methyl and other stable ethers are used in proving structures of glycosides. S-Adenosyl methionine is utilized for the methylation of many biologically important compounds such as tRNA in which some of the bases and some of the riboses are methylated.

C. Acetals and ketals

General reaction: Pair of hydroxyl groups + aldehyde/ketone ⟹ cyclic acetal/ketal

Acetone and other ketones react with sulfuric acid or copper sulphate, and zinc chloride to produce cyclic ketals usually with five-membered dioxolane ring systems. Aldehyde (Acetaldehyde, Formaldehyde, Benzaldehyde) gives cyclic acetals. The reaction is reversible and the groups can be removed by acid hydrolysis.

D. Glycoside formation

Reaction: Glycoside formation

This reaction takes place with excess ROH and acid catalyst. Alternatively, the "ROH" can be from another saccharide so joining two saccharide units. Glycosides are just acetals or ketals. The chemistry and uses are similar to just that of simple acetals and ketals. In case of a disaccharide formation the uses are indicated below.

E. Oxidation

General reaction: Alcohol ⟹ aldehyde, ketone or acid

Reagents DMSO + dicyclohexylcarbodiimide can oxidize isolated secondary hydroxyl groups to ketone and isolated primary hydroxyl groups to aldehyde. DMSO + acetic anhydride or phosphrous oxychloride are capable of similar oxidation.

Analogous biochemical process: Many sugar, alcohols and cyclitols are oxidized by enzyme utilizing the nicotinamide dinucleotides DPN^+ and TPN^+ .They are usually fairly specific and do not oxidize all alditols.

Oxidized sugras are used for the synthesis of rare sugars.

F. Ring Closure

In this reaction the aldehyde or ketone of the sugar reacts with an alcohol (hydroxyl) at the other end of the sugar. The alcohol group that reacts is the next-to-last one at the other end

of the chain. By tradition, the end alcohol group and the carbon it is bonded to are directed up. The oxygen of the next-to-last hydroxyl group becomes the connecting oxygen of the ether. The hydrogen of this group bonds with the alcohol of the aldehyde or ketone, creating another hydroxyl group. This hydroxyl group can be oriented either up or down. The alcohol groups in the middle should remain on the same side of the carbon as in the linear version of the structure (Fig. 14).

Fig. 14: Ring closure

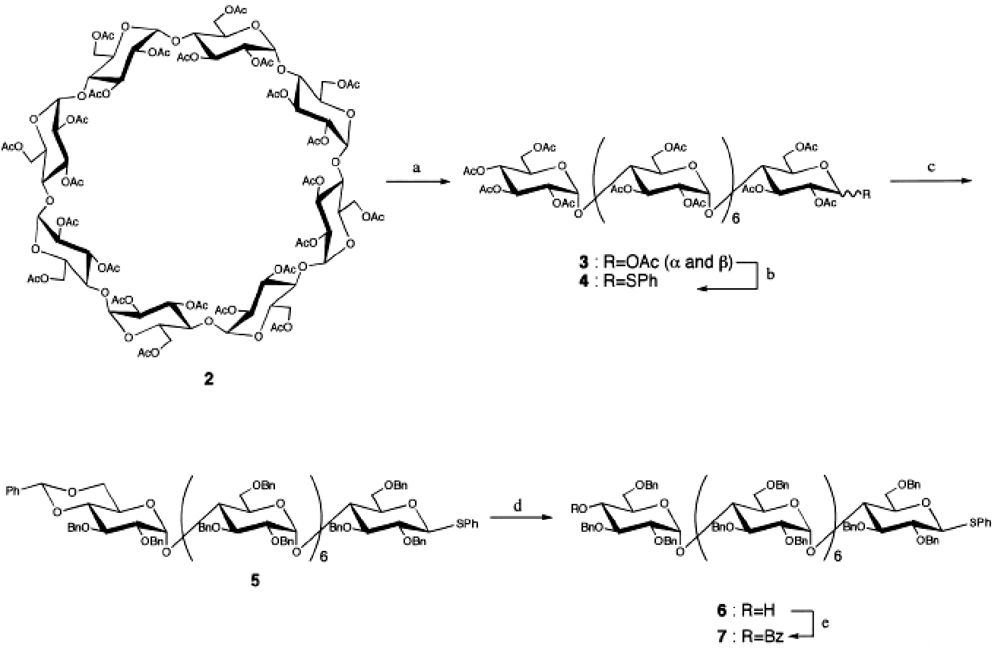
G. Polymerization

a) Chemical polymerization

1) The alcohol part of the hemiacetal (in the ring form of the sugar) reacts with an alcohol in another sugar (also in the ring form). For eg. , reaction of glucose & fructose leads to the production of sucrose.

D Glucose D Fructose Sucrose

2) Polycondensation and subsequent deprotection of a partially benzylated phenyl 1-thio-β-malto octaoside having a sole hydroxyl group at the non-reducing end gives rise to an amylase like α-(1,4)-glucan (Fig. 15).



Scheme 1. Chemical synthesis of α-(1,4)-glucan from γ-cyclodextrin. Reagents and conditions: (a) Acetic anhydride-conc. Sulfuric acid (50:1 v/v), 50−60°C, 24 h; (b) $PhSSiMe_3$-ZnI_2, 1,2-dichloroethane, RT, 40 h; (c) NaOMe/MeOH; $PhCH(OMe)_2$/CSA, DMF, 60°C,15 h; BnBr/NaH, DMF, RT, 20 h; (d) BH_3?Me_3N-$AlCl_3$, THF, 2 d; (e) BzCl, pridine, RT.

Fig. 15: Chemical polymerization

b) Enzymatic polymerization

Glycosyltransferases (EC 2.4) act as a catalyst for the transfer of a monosaccharide unit from an activated sugar phosphate (known as the "glycosyl donor") to an acceptor molecule, usually an alcohol. The result of glycosyl transfer can be a monosaccharide glycoside, an oligosaccharide, or a polysaccharide, although some glycosyltransferases catalyse transfer to inorganic phosphate or water. Glycosyl transfer can also occur to protein residues, usually to tyrosine, serine or threonine to give O-linked glycoproteins, or to asparagine to give N-linked glycoproteins. Mannosyl groups may be transferred to tryptophan to generate C-mannosyl tryptophan, which is relatively abundant in eukaryotes.

Commonly, sugar nucleotide derivatives are used as glycosyl donors. However, certain other glycosyltransferases utilize non-nucleotide donors, which may be polyprenol pyrophosphates, polyprenol phosphates, sugar-1-phosphates or sugar-1-pyrophosphates. Common examples include phosphorylase, glycogen synthase, chitinase, etc. (Fig. 16).

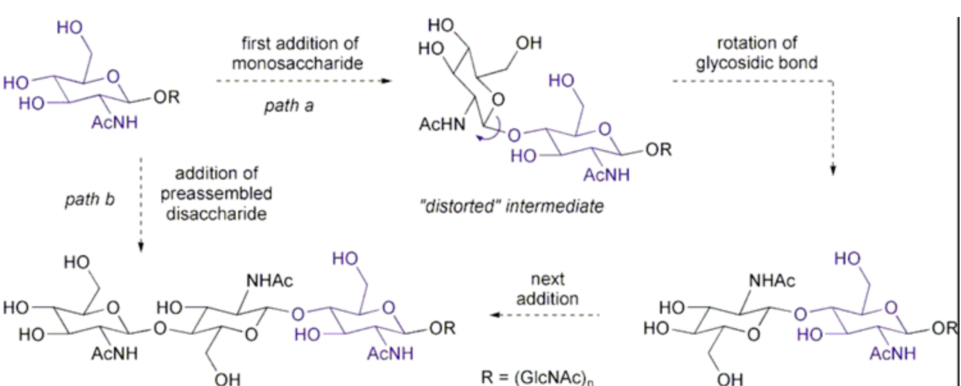


Fig. 16: Enzymatic polymerization

Monosaccharide, Disaccharides, Trisaccharide and Polysaccharide (Structure, occurrence and biological importance)

A. Monosaccharide

a. Glucose

Systemic name: ***D-Glucose***
Molecular formula: $\mathbf{C_6H_{12}O_6}$

Glucose (Glc), a monosaccharide (or simple sugar), is an important carbohydrate in biology. The living cell uses it as a source of energy and metabolic intermediate. Glucose is one of the main products of photosynthesis and starts cellular respiration in both prokaryotes and eukaryotes. The name comes from the Greek word 'glykys', which means "sweet", plus the suffix "-ose" which denotes a sugar.

Two stereoisomers of the aldohexose sugars are known as glucose, only one of which (D-glucose) is biologically active (Figs 2-4). This form (D-glucose) is often referred to as dextrose monohydrate, or, especially in the food industry, simply dextrose (from dextrorotatory glucose). The mirror-image of the molecule, L-glucose, cannot be metabolized by cells in the biochemical process known as glycolysis.

Function: Enzyme-regulated addition of glucose to proteins by glycosylation is often essential to their function.

- As an energy source - Glucose is a ubiquitous fuel in biology. It is used as an energy source in most organisms, from bacteria to humans. Use of glucose may be by either aerobic or anaerobic respiration (fermentation). Carbohydrates are the human body's key source of energy, through aerobic respiration, providing approximately 3.75 kilocalories (16 kilojoules) of food energy per gram. Breakdown of carbohydrates (e.g. starch) yields mono- and disaccharides, most of which is glucose. Through glycolysis and later in the reactions of the Citric acid cycle (TCA), glucose is oxidized to eventually form CO2 and water, yielding energy, mostly in the form of ATP. The insulin reaction and other mechanisms, regulate the concentration of glucose in the blood.
- Use of glucose as an energy source in cells is via aerobic or anaerobic respiration. Both of them start with the early steps of the glycolysis metabolic pathway. The first step of this is the phosphorylation of glucose by hexokinase to prepare it for later breakdown to provide energy.

As a precursor: Glucose is used as a precursor for the synthesis of several important substances. Starch solution Starch, cellulose, and glycogen are common glucose polymers (polysaccharides). Lactose, the predominant sugar in milk, is a glucose-galactose disaccharide. In sucrose, another important disaccharide, glucose is joined to fructose.

b. Galactose

Systemic name: ***D-Galactose***
Molecular formula: $\mathbf{C_6H_{12}O_6}$

Galactose (Gal) is a type of sugar, (Fig. 4) which is less sweet than glucose and not very water-soluble. Galactose is more commonly found in the disaccharide, lactose or milk

sugar. It is found as the monosaccharide in peas. Galactose is classified as a monosaccharide, an aldose, a hexose, and is a reducing sugar. It is considered a nutritive sweetener because it has food energy. A genetic defect of not being able to utilize galactose is called Galactosemia. The disorder is caused by a deficiency in one or more enzymes required to metabolize galactose. Since galactose is in milk as part of lactose, it builds up in the blood and urine. Undiagnosed it may lead to mental retardation, failure to grow, formation of cataracts, and in severe cases death by liver damage.

c. Fructose

Systemic name: ***D-Fructose***
Molecular formula: $\mathbf{C_6H_{12}O_6}$

Fructose or laevulose) is a simple reducing sugar (monosaccharide) (Fig. 3) found in many foods and is one of the three most important blood sugars along with glucose and galactose. Honey, tree fruits, berries, melons, and some root vegetables, such as beets, sweet potatoes, parsnips, and onions, contain fructose, usually in combination with sucrose and glucose. Fructose is also derived from the digestion of sucrose, a disaccharide consisting of glucose and fructose that is broken down by glycoside hydrolase enzymes during digestion. Fructose is the sweetest naturally occurring sugar, estimated to be twice as sweet as sucrose.

Fructose is often recommended for, and consumed by, people with diabetes mellitus or hyperglycemia, because it has a very low glycemic index (GI) relative to cane sugar (sucrose). The low GI is due to the unique and lengthy metabolic pathway of fructose, which involves phosphorylation and a multi-step enzymatic process in the liver.

d. 2 Deoxy-D-glucose:

Systemic name: ***2 Deoxy-D-arabino-hexose***
Molecular formula: $\mathbf{C_6H_{12}O_5}$

2-Deoxy-D-glucose is a glucose molecule (Fig. 17), which has the 2-hydroxyl group, replaced by hydrogen, so that it cannot undergo further glycolysis. It is also called as non-metabolizable form of glucose. The sugar is responsible for inhibiting glycosylating of proteins if used in culture medium.

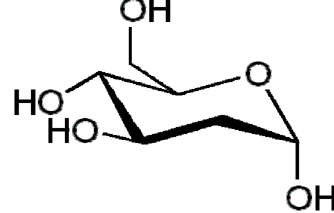


Fig. 17: 2 Deoxy-D-glucose

e. Xylose

Systemic name: ***D-Xylose***
Molecular formula: $\mathbf{C_5H_{10}O_5}$

Xylose or wood sugar is an aldopentose — a monosaccharide containing five carbon atoms and including an aldehyde functional group (Fig. 4). It has chemical formula $C_5H_{10}O_5$. Xylose is found in the embryos of most edible plants.

Function

- Xylose is an antibacterial and antifungal, particularly with gram-negative organisms and *Candida* sp.
- Absorption rate of Xylose is decreased in some patients with intestinal disorders, including those with colitis and diabetes, suggesting that when absorption problems are corrected, these conditions might also reverse themselves.
- Unlike sucrose or artificial sweeteners, xylose promotes the growth of "friendly flora" in the intestines, thus increasing the manufacture and absorption of all nutrients and strengthening the immune system to help fight off any type of illness.

f. Ribose

Systemic name: ***D-ribose***
Molecular formula: $\mathbf{C_5H_{10}O_5}$

Ribose, primarily seen as D-ribose, is an aldopentose — a monosaccharide containing five carbon atoms, and including an aldehyde functional group in its linear form. It has the chemical formula $C_5H_{10}O_5$, and was discovered in 1905 by Phoebus Levene (Fig. 18).

Fig. 18: Ribose

As a component of the RNA that is used for genetic transcription, ribose is critical to living creatures. It is related to deoxyribose, which is a component of DNA. It is also a component of ATP, NADH, and several other chemicals that are critical to metabolism.

g. Deoxyribose

Systemic name: ***2 -Deoxy-D-erythro-pentoseD-ribose***
Molecular formula: $\mathbf{C_5H_{10}O_4}$

Deoxyribose, also known as D-Deoxyribose and 2-deoxyribose, is an aldopentose — a monosaccharide containing five carbon atoms, and including an aldehyde functional group in its linear structure (Fig. 19). It is a deoxy sugar derived from the pentose sugar ribose by the replacement of the hydroxyl group at the 2 position with hydrogen, leading to the net loss of an oxygen atom.

Ribose forms a five-member ring composed of four carbon atoms and one oxygen. Hydroxyl groups are attached to three of the carbons. The other carbon and a hydroxyl group are attached to one of the carbon atoms adjacent to the oxygen. In deoxyribose, the carbon furthest from the attached carbon is stripped of the oxygen atom in what would be a

hydroxyl group in ribose. Due to the common C3' and C4' stereochemistry of D-ribose and D-arabinose, D-2-deoxyribose is also D-2-deoxyarabinose.

Fig. 19: Deoxyribose

Biological importance

Ribose and 2-deoxyribose derivatives have an important role in biology. Among the most important derivatives are those with phosphate groups attached at the 5 position. Mono-, di-, and triphosphate forms are important, as well as 3-5 cyclic monophosphates. Purines and pyrimidines form an important class of compounds with ribose and deoxyribose. When these purine and pyrimidine derivatives are coupled to a ribose sugar, they are called nucleosides. In these compounds, the convention is to put a ′ (pronounced "prime") after the carbon numbers of the sugar, so that in nucleoside derivatives a name might include, for instance, the term "5′-monophosphate", meaning that the phosphate group is attached to the fifth carbon of the sugar, and not to the base. The bases are attached to the 1′ ribose carbon in the common nucleosides. Phosphorylated nucleosides are called nucleotides.

One of the common bases is adenine (a purine derivative); coupled to ribose it is called adenosine; coupled to deoxyribose it is called deoxyadenosine. The 5′-triphosphate derivative of adenosine, commonly called ATP, for adenosine triphosphate, is an important energy transport molecule in cells.

2-Deoxyribose and ribose nucleotides are often found in unbranched 5′-3′ polymers. In these structures, the 3′carbon of one monomer unit is linked to a phosphate that is attached to the 5′carbon of the next unit, and so on. These polymer chains often contain many millions of monomer units. Since long polymers have physical properties distinctly different from those of small molecules, they are called macromolecules. The sugar-phosphate-sugar chain is called the backbone of the polymer. One end of the backbone has a free 5′phosphate, and the other end has a free 3′OH group. The backbone structure is independent of which particular bases are attached to the individual sugars.

Genetic material in earthly life often contains poly 5′-3′, 2′-deoxyribose nucleotides, in structures called chromosomes, where each monomer is one of the nucleotides deoxy-adenine, thymine, guanine or cytosine. This material is commonly called deoxyribonucleic acid, or simply DNA for short. DNA in chromosomes forms very long helical structures containing two molecules with the backbones running in opposite directions on the outside of the helix and held together by hydrogen bonds between complementary nucleotide bases lying between the helical backbones. The lack of the 2′ hydroxyl group in DNA appears to allow the backbone the flexibility to assume the full conformation of the long double-helix, which involves not only the basic helix, but additional coiling necessary to fit these very long molecules into the very small volume of a cell nucleus.

In contrast, very similar molecules, containing ribose instead of deoxyribose, and known generically as RNA, are known to form only relatively short double-helical complementary base paired structures. These are well known, for instance, in ribosomal RNA molecules and in transfer RNA (tRNA), where so-called hairpin structures from palindromic sequences within one molecule.

B. Disaccharide

a. Sucrose

Systematic name: ***D-glucopyranosyl-(1↔2)-β-D-fructofuranoside***
Molecular formula: $\mathbf{C_{12}H_{22}O_{11}}$

Sucrose **(Fig. 5)** (common name: table sugar, also called saccharose) is a disaccharide (glucose + fructose) best known for its role in human nutrition and is formed by plants but not by other organisms such as animals. Sucrose is an easily assimilated macronutrient that provides a quick source of energy to the body, provoking a rapid rise in blood glucose upon ingestion. However, pure sucrose is not normally part of a human diet balanced for good nutrition, although it may be included sparingly to make certain foods more palatable.

Sucrose, as a pure carbohydrate, has an energy content of 3.94 kilocalories per gram (or 17 kilojoules per gram). When a large amount of foods that contain a high percentage of sucrose is consumed, beneficial nutrients can be displaced from the diet, which can contribute to an increased risk for chronic disease such as defect in glucose metabolism or diabetes mettlitus

b. Cellobiose

Systemic name: *4-O- β-D-glucopyranosyl-D-glucose*
Molecular formula $C_{12}H_{22}O_{11}$

Cellobiose is a disaccharide derived from the condensation of two glucose molecules linked in a β(1→4) bond. It can be hydrolyzed by bacteria or cationic ion exchange resins to give glucose. Cellobiose has eight free alcohol (COH) groups and three ether linkages, which give rise to strong inter- and intra-molecular hydrogen bonds.

c. Trehalose

Systemic name: *α-D-glucopyranosyl- α-D-glucopyranoside*
Molecular formula: $C_{12}H_{22}O_{11}$

Trehalose, also known as mycose, is a natural alpha-linked disaccharide formed by an α, α-1, 1-glucoside bond between two α-glucose units (Fig. 5). It can be synthesised by fungi, plants, and invertebrate animals. It is implicated in anhydrobiosis — the ability of plants and animals to withstand prolonged periods of desiccation. It has high water retention capabilities and is used in food and cosmetics. The sugar forms a gel phase as cells dehydrate, which prevents disruption of internal cell organelles by effectively splinting them in position. Rehydration then allows normal cellular activity to be resumed without the major, lethal damage that would normally follow a dehydration/reyhdration cycle. Trehalose has the added advantage of being an antioxidant. Trehalose can be found in nature, animals, plants, and microorganisms. In animals, trehalose is prevalent in shrimp, and also in insects, including grasshoppers, locusts, butterflies, and bees, in which blood-sugar is trehalose. The trehalose is then broken down into glucose by the catabolic enzyme

trehalase for use. In plants, the presence of trehalose is seen in sunflower seeds, selaginella mosses, and sea algae. Within the fungus family, it is prevalent in mushrooms shiitake (*Lentinula edodes*), maitake (*Grifola fondosa*), nameko (*Pholiota nameko*), and Judas's ear (*Auricularia auricula-judae*) contain 1% to 17% percent of trehalose in dry weight form. Trehalose is found in such microorganisms as baker's yeast and wine yeast. When tardigrades (water bears) dry out, the glucose in their bodies changes to trehalose when they enter a state called cryptobiosis - a state wherein they appear dead. However, when they receive water, they revive and return to their metabolic state. Larva of sleeping chironomid (polypedihum vanderplanki) and artemia (sea monkeys, brine shrimp) can withstand dehydration, because they store trehalose within their cells. In the plant kingdom, selaginella mosses that grow in desert and mountainous areas, although they may be cracked and dried out, will turn green again and revive after a rain, because of the function of Trehalose. Dried shitake mushrooms spring back into shape in water because they contain trehalose.

Trehalose is metabolized by a number of bacteria, including *Streptococcus mutans*, the common oral bacteria responsible for dental plaque. The enzyme trehalase, a glycoside hydrolase, breaks trehalose into two glucose molecules, which can then be readily absorbed in the gut. Trehalose is the major carbohydrate energy storage molecule used by insects for flight to be used as glucose for the rapid energy requirement of flight.

d. Lactose

Systemic name: ***4-O- β-D-galactopyranosyl-D-glucose***
Molecular formula: $\mathbf{C_{12}H_{22}O_{11}}$

Lactose (also referred to as milk sugar) is a sugar, which is found most notably in milk (Fig. 5). Lactose makes up around 2–8% of milk (by weight). The name comes from the Latin word for milk, plus the -ose ending used to name sugars. Lactose is the only significant sugar or carbohydrate of animal origin. Lactose is a disaccharide that consists of β-D-galactose and β-D-glucose fragments bonded through a β(1→4) glycosidic linkage.

An enzyme, lactase, is essential for digestion of lactose, and a majority of adults in this country do not have lactase, which digests milk sugar, or lactose. Thus, it is very difficult for them to digest milk and may contribute to gaseousness, cramping, and diarrhea.

e. Maltose

Systemic name: ***4-O-α -D-glucopyranosyl-D-glucose***
Molecular formula: $\mathbf{C_{12}H_{22}O_{11}}$

Maltose, or malt sugar, is a disaccharide formed from two units of glucose joined with an α(1→4) linkage. It is the second member of an important biochemical series of glucose chains. The addition of another glucose unit yields maltotriose; further additions will produce dextrins (also called maltodextrins) and eventually starch.

Maltose can be broken down into two glucose molecules by hydrolysis. In living organisms, the enzyme maltase can achieve this very rapidly. In the laboratory, heating with a strong acid for several minutes will produce the same result. The production of maltose from germinating cereals, such as barley, is an important part of the brewing process. When barley is malted, it is brought into a condition in which the concentration of maltose-producing amylases has been maximized. Mashing is the process by which these amylases

convert the cereal's starches into maltose. Metabolism of maltose by yeast during fermentation then leads to the production of ethanol and carbon dioxide.

C. Trisaccharides

a. Raffinose

Systemic name: ***β-D-Fructofuranosyl-O-α-D-galactopyrnosyl-(1→6)- α-D-glucopyranoside***
Molecular formula: $\mathbf{C_{18}H_{32}O_{11}}$

Raffinose is a trisaccharide composed of galactose, fructose, and glucose (Figure 20). It can be found in beans, cabbage, brussels sprouts, broccoli, asparagus, other vegetables, and whole grains. Raffinose can be hydrolyzed to D-galactose and sucrose by the enzyme α-galactosidase (α-GAL), an enzyme not found in humans. α-GAL also hydrolyzes other α-galactosides such as stachyose, verbascose, and galactinol, if present. The enzyme does not cleave β-linked galactose, as in lactose. The raffinose families of oligosaccharides (RFOs) are alpha-galactosyl derivatives of sucrose, and the most common are the trisaccharide raffinose, the tetrasaccharide stachyose, and the pentasaccharide verbascose. RFOs are almost ubiquitous in the plant kingdom, being found in a large variety of seeds from many different families, and they rank second only to sucrose in abundance as soluble carbohydrates. Raffinose is also used as a base substance for sucralose.

Fig. 20: Raffinose

Humans and other monogastric animals (pigs and poultry) do not possess the α-GAL enzyme to break down RFOs and these oligosaccharides pass undigested through the stomach and upper intestine. In the lower intestine, they are fermented by gas-producing bacteria which do posses the α-GAL enzyme and make carbon dioxide, methane, and/or hydrogen -- leading to the flatulence commonly associated with eating beans and other vegetables. α-GAL is present in digestive aids such as the product Beano.

b. Acarbose

Systemic name: ***O-4,6-Dideoxy-4[[[1S-(1α,4α,5β,6α)]-4,5,6-trihydroxy-3-(hydroxymethyl)-2-cyclohexen-1yl] amino]-α-D-glucopyranosyl-(1→4)-O- α-D-glucopyranosyl-(1→4)-D-glucose***
Molecular formula: $\mathbf{C_{25}H_{43}NO_{18}}$

Acarbose is an anti-diabetic drug used to treat type 2 diabetes mellitus and, in some countries, prediabetes. (Fig. 21) It is an inhibitor of alpha glucosidase, an enteric enzyme that releases glucose from larger carbohydrates.

Fig. 21: Acarbose

c. Melezitose

Systemic name*: **O- α -D-glucopyranosyl-(1→3)- β-D- fructofuranosyl- α-D-glucopyranoside***
Molecular formula: $\mathbf{C_{18}H_{32}O_{16}}$

Melezitose, also spelled melicitose, is a nonreducing trisaccharide sugar (Fig. 22) that is produced from lice such as *Cinara pilicornis* by an enzyme reaction. This is beneficial to the insects, as it reduces the stress of osmosis by reducing their own water potential. The melezitose is part of the honeydew of the honey produced by bees, and acts as an attractant for ants. This is useful to the lice as they have a symbiotic relationship with ants. Melezitose can be partially hydrolyzed to glucose and turanose, the latter of which is an isomer of sucrose.

Since acarbose prevents the degradation of complex carbohydrates into glucose, the carbohydrates will remain in the intestine. In the colon, bacteria will digest the complex carbohydrates, thereby causing gastrointestinal side effects such as flatulence (78% of patients) and diarrhea (14% of patients).

Fig. 22: Melezitose

D. Polysaccharides

Carbohydrate composed of ten or more monosaccharide units joined together by glycosidic linkages are classified as polysaccharides. Most of the carbohydrates found in nature occur as polysaccharides, polymers of high molecular weight; which on hydrolysis yield monosaccharides or products related to monosaccharides, most frequently D-glucose. Others monosaccharides obtained are D-mannose. D- and L-galactose, L-arabinose, D-glucuronic acid, D- and L-galactose, L-arabinose. D-glucuronic acid, D-glucosamine, etc. 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. Polysaccharides can be classified-as homopolysaccharidcs and heleropolysciccharide (Fig. 23) depending on the variety of sugar moieties joined to the sugar chain. Some are linear polymers and others are branched. Some typical polysaccharides are starches, glycogens celluloses, agar, gum chitin, etc.

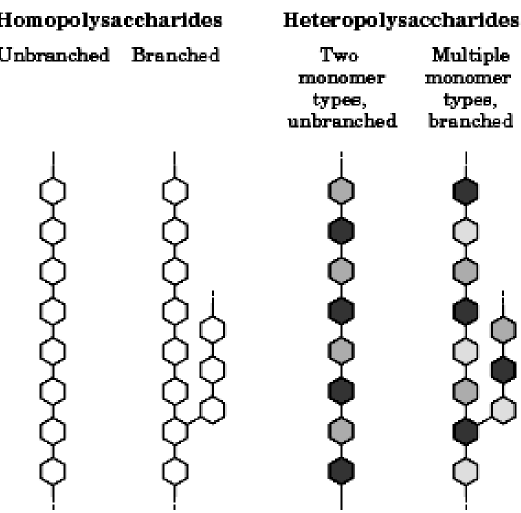


Fig. 23: Polysaccharide

Some homopolysaccharides serve as storage forms of monosaccharides used as fuels; starch and glycogen are homopolysaccharides of this type. Other homopolysaccharides (cellulose and chitin) serve as structural elements in plant cell walls and animal exoskeletons. Heteropolysaccharides provide extracellular support for organisms of all kingdoms. The rigid layer of the bacterial cell envelope (the peptidoglycan) is a heteropolysaccharide built from two alternating monosaccharide units. In animal tissues, several heteropolysaccharides, occupy the extracellular space forming a matrix that holds individual cells together and provides protection, shape, and support to cells, tissues, or organs. Hyaluronic acid accounts for the toughness and flexibility of cartilage and tendon, is among this group of extracellular polysaccharides. Other heteropolysaccharides, sometimes in very large aggregates with proteins (proteoglycans), account for the high viscosity and lubricating properties of some extracellular secretions.

Unlike proteins, polysaccharides generally do not have definite molecular weights, due to consequence of the mechanisms of assembly of the two types of polymers. Proteins are synthesized on a template (messenger RNA) of defined sequence and length, by enzymes that copy the template exactly. For polysaccharide synthesis, with no template; the program for polysaccharide synthesis is intrinsic to the enzymes that catalyze the polymerization of monomeric units. For each type of monosaccharide to be added to the growing polymer there is a separate enzyme, which acts only when the enzyme that inserts the preceding subunit has acted. The alternating action of several enzymes produces a polymer with a precise repeating sequence, but the exact length varies from molecule to molecule, within a general size class.

1. Homopolysaccharides

A homopolysaccharide is made up of a single kind of mono-saccharide. Starches, glycogens, and celluloses, which are made up of only glucose residues, are example of homopolysaccharides.

i) Starch

The starches occur widespread as reserve carbohydrate in tubers such as potatoes, in many fruits, grains and seeds. In the grains, the starch is arranged in concentric layers. When starch grains are treated with boiling water, the substance in the center passes into solution, but the greater part of the grain is not soluble. This insoluble portion swells as it absorbs water and the whole mass becomes starch paste. Both the soluble portion and the insoluble portion are heterogeneous mixtures. The soluble fraction is referred to as amylose, and the insoluble fraction as amylopectin. Most starches contain 80-90 per cent amylopectin and 10—20 per cent amylose. Amylose and amylopectin can be separated by taking advantage of the difference in solubility in water. Both amylose and amylopectin are polymers of glucose and upon hydrolysis with acid they give D-glucose as the product.

Structure of amylose: Amylose is an unbranched long chain polymer in which the glucose residues are linked through α-1, 4- glycosidic linkages. The amylose structure may be ragarded as a repeated maltose structure with a free sugar group (acetal group) at one end. This is also known as the reducing end (nth residue), whereas, the opposite end (first residue) is referred to as the non-reducing end. Any particular preparation of amylose usually consists of a mixture of populations of molecules, which differ widely in chain length (number of glucose residues per chain) (Fig. 24).

Fig. 24: Amylose

Structure of amylopectin: Amylopectin is also made of chains of glucose residues, but the chains are highly branched. The glucose residue situated at the branching point is substituted not only on carbon 4 but also on carbon 6 (Fig. 25). Both amylose and amylopectin give characteristic colour reactions with iodine. Amylose produces a blue-black colour, whereas, amylopectin gives a purple colour.

Saccharification of starch by amylases: Amylases are enzymes of plant and animal origin and they hydrolyse starch. There are two kinds of amylases, alpha-amylases and beta-amylases. Alpha-amylases are endo enzymes, act on amylose and amylopectin in a random fashion. Initially more central linkages are cleaved and smaller polysaccharide chains are formed. As the reaction proceeds, these are further hydrolysed to maltose and glucose.

α-amylases hydrolyse amylose and amylopectin in an orderly fashion. It acts from one end of the polymer and cleaves off two glucose residues at a time, as a maltose unit. As the reaction proceeds, amylose is completely hydrolysed to maltose. However, with the

amylopectin, the reaction stops as a branch point is approached. The terminal parts of the branches are digested away as maltose units in an orderly fashion and the central core is left behind as the enzyme is blocked at 1-6 glycosidic linkages or branch points. The polysaccharide fragment that remains after such incomplete hydrolysis is called a dextrin.

Fig. 25: Amylopectin

Dextrins formed from amylopectin by β-amylase are highly branched and give a red colour with iodine. Dextrins of relatively small molecular size do not give a colour with iodine. The colour produced by reaction with iodine is used as an indication of the degree of branching of starch. Starch is readily hydrolysed by dilute mineral acid with ultimate formation of glucose in quantitative yield. The course of hydrolysis may be followed by the gradual change in colour produced by iodine: blue-black-purple-red-colourless.

ii) Celluloses

Cellulose is the most abundant organic compound in nature and it is the chief constituent of the fibrous parts of plants. The purest form of cellulose is usually obtained from cotton. Cellulose is a polymer made up of glucose residues. Upon hydrolysis, cellulose yields D-glucose as the product. Cellulose molecules are not branched and consist essentially of long chains with glucose residues linked in repeating sequence of cellobiose structures. The glucose residues are linked together through β-l, 4-glycosidic linkages. The celluloses obtained from different sources differ in molecular size though they are all made up of glucose. Cellulose is found in plants as microfibrils (2-20 nm diameter and 100 - 40 000 nm long). These form the structurally strong framework in the cell walls.

***Structural unit*:** Cellulose is a linear polymer of β-(1→4)-D-glucopyranose units in ${}^4C_1$ conformation (Fig. 26). The fully equatorial conformation of β-linked glucopyranose residues stabilizes the chair structure, minimizing its flexibility (for example, relative to the slightly more flexible α-linked glucopyranose residues in amylose). Cellulose preparations may contain trace amounts (~0.3%) of arabinoxylans

***Molecular structure*:** Cellulose is an insoluble molecule consisting of between 2000 - 14000 residues with some preparations being somewhat shorter. It forms crystals (cellulose Iα) where intra-molecular (O3-H→O5' and O6→H-O2') and intra-strand (O6-H→O3') hydrogen bonds holds the network flat allowing the more hydrophobic ribbon faces to stack. Each residue is oriented 180° to the next. Although individual strand of cellulose are intrinsically no less hydrophilic, or no more hydrophobic, than some other soluble polysaccharides (such as amylose) this tendency to form crystals utilizing extensive intra-

and intermolecular hydrogen bonding makes it completely insoluble in normal aqueous solutions. It is thought that water molecules catalyze the formation of the natural cellulose crystals by helping to align the chains through hydrogen-bonded bridging. Part of a cellulose preparation is amorphous between these crystalline sections. The overall structure is of aggregated particles with extensive pores capable of holding relatively large amounts of water by capillarity. The natural crystal is made up from metastable Cellulose I with all the cellulose strands parallel and no inter-sheet hydrogen bonding. This cellulose I (natural cellulose) contains two coexisting phases cellulose Iα (triclinic) and cellulose Iβ (monoclinic) in varying proportions dependent on its origin (Fig. 27).

Fig. 26: Cellulose

Function – Cellulose is the structural component of the primary cell wall of green plants, many forms of algae and the oomycetes. Some species of bacteria secrete it to form biofilms.

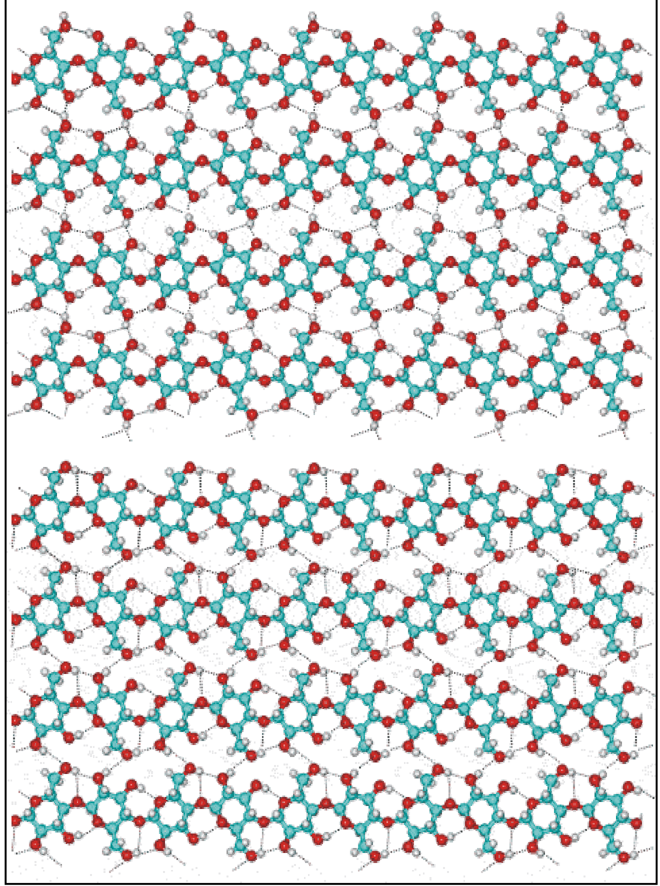


Fig. 27: Organization of glucose chains in cellulose

iii) Glycogen

Glycogen is the reserve carbohydrate in the animals and it is found in significant amounts in the liver and muscle. Glycogen is made up of D-glucose residues. Upon hydrolysis, it yields D-glucose as the product. Glycogen is a highly branched chain polysaccharide and it resembles amylopectin in structure. Thus glucose residues are linked together through α-l,4-glycosidic linkages except at the branch points. The branch is linked to the main chain through α-l,6-glycosidic linkages. It is very similar to the structure of amylopectin. The average molecular weights of glycogen preparations vary from 270,000 to 100,000,000. A single preparation usually consists of a mixture of populations of molecules that differ in molecular weight (Fig. 28). In the structure of glycogen, O represents glucose residues. The arrows indicate the α-1→6 glycosidic linkage. All the other glycosidic linkage are α-1,4 . The asterisks indicate the non—reducing end groups and r.e. strands for reducing end group.

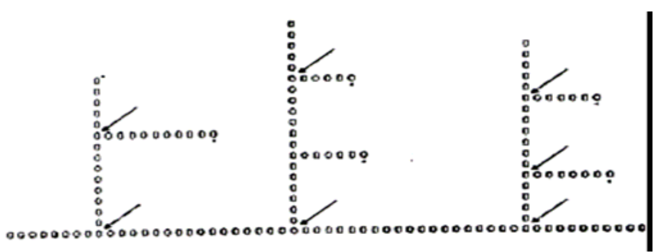


Fig. 28: Glycogen

In a wide range of organisms, excess glucose is stored not as monomer but converted to polymeric forms for storage—glycogen in vertebrates and many microorganisms, starch in plants. With two minor differences, glycogen has almost the same structure as amylopectin (a constituent of starch). The glycogen molecule is roughly twice as large as amylopectin, and branching frequency—that also controls the mobilization of stored polysaccharide. Glycogen has roughly twice as many branches (branching occurs on an average after every 8-12 residues unlike amylopectin where it is after every 24-30 residues). There is an advantage to branched polysaccharides such as amylopectin and glycogen. During times of shortage, enzymes attack one end of the polymer chain and cut off glucose molecules, one at a time. More the branches, more the points at which the enzyme attacks the polysaccharide. Thus, a highly branched polysaccharide is better suited for the rapid release of glucose than a linear polymer.

In vertebrates, glycogen is found primarily in the liver and skeletal muscle but can also be made by the brain, uterus, and the vagina. However, muscle glycogen is not generally available to other tissues, because muscle lacks the enzyme glucose-6-phosphatase. Glycogen may represent up to 10% of the weight of liver and 1% to 2% of the weight of muscle. Stores of glycogen in the liver are considered the main buffer of blood glucose levels. If this much glucose was dissolved in the cytosol of a hepatocyte, its concentration would be about 0.4 M, enough to dominate the osmotic properties of the cell. When stored as a long polymer (glycogen), however, the same mass of glucose has a concentration of only 0.01 μM. Glycogen is stored in large cytosolic granules. The elementary particle of glycogen, the β particle, about 21 nm in diameter, consists of up to 55,000 glucose residues with about 2,000 nonreducing ends. The major site of daily glucose consumption (75%) is the brain via aerobic pathways. Most of the remainder of it is utilized by erythrocytes, skeletal muscle, and heart muscle. Due to disbalance in glucose utilization iabetics may experience

Glycogen degradation and synthesis are relatively simple biochemical processes. Glycogen synthesis differs from glycogen breakdown. Unlike breakdown, synthesis is endergonic, meaning that glycogen is not synthesized without the input of energy. Energy for glycogen synthesis comes from UTP, which reacts with glucose-1-phosphate, forming UDP-glucose, in reaction catalysed by UDP-glucose pyrophosphorylase. Glycogen is synthesized from monomers of UDP-glucose by the enzyme Glycogen synthase, which progressively lengthens the glycogen chain. As glycogen synthase can only lengthen an existing chain, the protein glycogenin is needed to initiate the synthesis of glycogen.

Glycogen degradation consists of three steps: (1) the release of glucose 1-phosphate from glycogen, (2) the remodeling of the glycogen substrate to permit further degradation, and (3) the conversion of glucose 1-phosphate into glucose 6-phosphate for further metabolism. It is cleaved from the nonreducing ends of the chain by the enzyme glycogen phosphorylase to produce monomers of glucose-1-phosphate that is then converted to glucose 6-phosphate. A special debranching enzyme is needed to remove the α (1-6) branches in branched glycogen and reshape the chain into linear polymer. Debranching enzyme has two independent active sites, consisting of residues in different segments of a single polypeptide chain, that catalyze α(1-6) glucosidase and transferase (transglycosylase) reactions.

The transferase of the debranching enzyme transfers three glucose residues from a 4-residue limit branch to the end of another branch, diminishing the limit branch to a single glucose residue. The α(1-6) glucosidase moiety of the debranching enzyme then catalyzes hydrolysis of the α(1-6) linkage, yielding free glucose. This is a minor fraction of glucose released from glycogen. The major product of glycogen breakdown is glucose-1-phosphate, arising from phosphorylase activity, which is subsequently converted, to glucose-6-phosphate. The G6P monomers produced have three possible fates:

- G6P can continue on the glycolysis pathway and be used as fuel.
- G6P can enter the pentose phosphate pathway via the enzyme glucose-6-phosphate dehydrogenase to produce NADPH and 5-carbon sugars.
- In the liver and kidney, G6P can be dephosphorylated back to glucose by the enzyme glucose 6-phosphatase. This is the final step in the gluconeogenesis pathway.

v) Chitin

Chitin is a linear homopolysacchande composed of N-acetyl- D-glucosamine residues in β-linkage. The only chemical difference from cellulose is the replacement of a hydroxyl group a: C-2 with an acetylated amino group. Chitin forms extended fibers similar to those of cellulose, and structure of glycogen. Chitin like cellulose is indigestible by vertebrate animals. It is the main component of the cell walls of fungi, the exoskeletons of arthropods, such as crustaceans (like the crab, lobster and shrimp) and the insects, including ants, beetles and butterflies, the radula of mollusks and the beaks of the cephalopods, including squid and octopuses. Chitin is probably the second most abundant polysaccharide next to cellulose, in nature (Fig. 29).

Chitin is used in water purification, and as an additive to thicken and stabilize foods and pharmaceuticals. It also acts as a binder in dyes, fabrics, and adhesives. Industrial separation membranes and ion-exchange resins can be made from chitin. Processes to size and strengthen paper employ chitin. Its properties as a flexible and strong material make it favorable as surgical thread. Its biodegradability means it wears away with time as the wound heals. Moreover, the polysaccharide has some unusual properties that accelerate

healing of wounds in humans. Most recent studies point out that chitin is a good inductor for defense mechanisms in plants. It is being tested as a fertilizer that can help plants develop healthy immune responses, and have a much better yield and life expectancy.

Fig. 29: Chitin

2: Heteropolysaccharides

A heteropolysaccharide is made up of two or more different monosaccharides, or monosaccharide derivatives.

i) Agar

Agar (agar-agar, kanten, agal-agal, ceylon agar, china grass) is a gelatinous substance chiefly used as a solid substrate to culture medium for microbiological work. It is an unbranched polysaccharide obtained from the cell membranes of some species of red algae or seaweed. Chemically, agar is a polymer made up of subunits of the sugar galactose. Agar polysaccharides serve as the primary structural support for the algae's cell walls. Agar is a heterogeneous mixture of two classes of polysaccharide: agaropectin and agarose. Although both polysaccharide classes share the same galactose-based backbone, agaropectin is heavily modified with acidic side-groups, such as sulfate and pyruvate (Fig. 30). Agarose is composed of agarobiose repeating disaccharide units alternating with 1, 3-linked-β-D-galactopyranose and 1, 4-linked-3, 6-anhydro-α-L-galactopyranose. The agaropectin seems to have the same backbone as the agarose, but contains considerable amount of acid groups such as sulfate, pyruvate and glucuronate groups.

Fig. 30: Agar

Property Agar is insoluble in cold water but dissolves to give random coils in boiling water. Gelation is reported to follow a phase separation process (although these findings are disputed) and association on cooling (~35°C), forming gels with up to 99.5% water and remaining solid up to about 85°C. It is derived primarily from *Gracilaria* species.

***Uses*:** Agar has a major use in microbiological media as it is not easy for microorganisms to metabolize and forms clear, stable and firm gels, but in the food area it is used in icings, glazes, processed cheese, jelly sweets and marshmallows. It may be used in tropical countries and by vegetarians as a substitute for gelatin.
Agar is 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.

ii) Agarose: 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. The two major components of agar **(Fig. 31)** are the unbranched polymer agarose (Mr ~120,000) (Figure 34) and a branched component, agaropectin. The remarkable gel-forming property of agarose makes it useful in the biochemistry. 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. Exterior hydroxyl groups allow aggregation of up to 10,000 of these helices to form suprafibers. Agarose gels are used as inert supports for the electrophoretic separation of nucleic acids, an essential part of the DNA sequencing process.

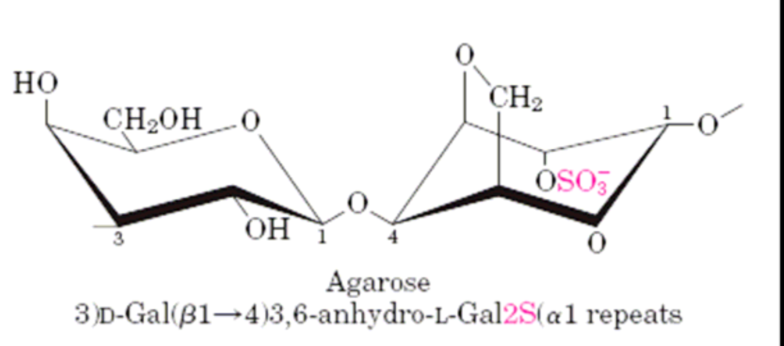


Fig. 31: Agarose

iii) Xylan

Lignocelluloses in plant kingdom are composed of cellulose, hemicellulose and lignin. 15-30% of the lignocellulose is hemicellulose, where xylan is the main component. They are almost as ubiquitous as cellulose in plant-cell walls and contain predominantly beta-D-xylose units linked as in cellulose. Xylan is a heteropolysaccharide found in plant cell walls and some algae (Fig. 32). It is found in almost all parts of the plant, in the cell walls of some green algae, especially macrophytic siphonous genera, where it replaces cellulose. Similarly, it replaces the inner fibrillar cell-wall layer of cellulose in some red algae. Xylans are polysaccarides mainly having backbone of 1,4 beta-linked xylose. The backbone is branched with sugars like glucose, arabinose, acetylated sugars and esters through alpha linkage. Xylans are of little commercial importance, but being associated with cellulose, forms a part of sugar hydrolysis project to be converted to alcohol and other value added products. In tune with the emergent need for production of ethanol from sugars xylan as hemicellulose forms the main component, which supports complete utilization of cellulose through enzymatic saccharification.

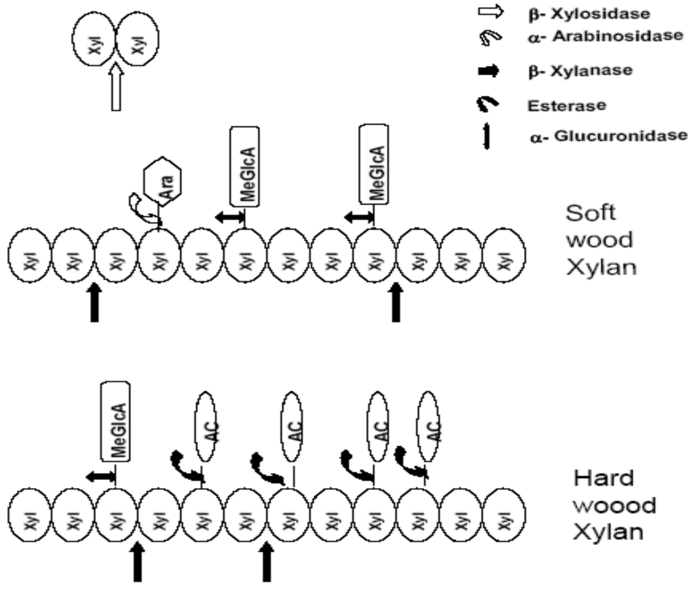


Fig. 32: Xylan

iv) Alginate

Alginates are produced by brown seaweeds (Phaeophyceae, mainly Laminaria). Alginates are linear unbranched polymers containing β-(1-4)-linked D-mannuronic acid (M) and α-(1-4)-linked L-guluronic acid (G) residues (Fig. 33). Although these residues are epimers (D-mannuronic acid residues being enzymatically converted to L-guluronic acid after polymerizationa) and only differ at C5, they possess very different conformations; D-mannuronic acid being 4C1 with diequatorial links between them and L-guluronic acid being 1C4 with diaxial links between them. Bacterial alginates are additionally O-acetylated on the 2 and/or 3 positions of the D-mannuronic acid residues. The bacterial O-acetylase may be used to O-acetylate the algal alginates, so increasing their water binding.

Fig. 33: Alginate

Molecular structure: Alginates are not random copolymers but, according to the source algae, consist of blocks of similar and strictly alternating residues (that is, MMMMMM, GGGGGG and GMGMGMGM), each of which have different conformational preferences and behavior (Fig. 33). As examples, the M/G ratio of alginate from *Macrocystis pyrifera* is about 1.6 whereas that from *Laminaria hyperborea* is about 0.45.

Functionality: The primary functions of the alginates are as thermally stable cold setting gelling agents in the presence of calcium ions; gelling at far lower concentrations than gelatin. Such gels can be heat treated without melting, although they may eventually degrade.

Alginate's solubility and water-holding capacity depend on pH (precipitating below about pH 3.5), molecular weight (lower molecular weight calcium alginate chains with less than 500 residues showing increasing water binding with increasing size), ionic strength (low ionic strength increasing the extended nature of the chains) and the nature of the ions present. Generally alginates show high water absorption and may be used as low viscosity emulsifiers and shear-thinning thickeners. They can be used to stabilize phase separation in low fat fat-substitutes.

Sugar derivatives (Structure, occurrence & functions)

Various sugar derivatives, which are essential participants in various reactions, are known and well characterized. Some of them participate in the transformations of simple sugars to other simple sugars or sugar derivatives. Here properties of some of the important derivatives of the sugars will be presented.

1. Derivatives of Monosaccharide

Monosaccharide undergoes a number of reactions to form biologically important derivatives. Three common types of monosaccharide derivatives are amino sugars, carboxylic acid sugars and sugar alcohols.

A. Amino sugars and N-acetylated sugars

An amino sugar contains an amine group in place of a hydroxyl group. Derivatives of amine containing sugars, such as N-acetylglucosamine and sialic acid, while not formally containing an amine, are also considered amino sugars. N-acetylated sugars are derivatives of amino sugars. The N in "N-acetyl" refers to the fact that the acetyl group is bonded to the nitrogen. The acetyl group here replaces one of the amine hydrogens. Connective tissue polysaccharides, such as cartilage, contain amino sugar and N-acetyl sugar monomers (Figs 34 and 35).

Significant amounts of glucosamine have been found in the intestinal mucin, which binds chloresterol, thereby limiting its absorption. Glucosamine has proven to decrease insulin secretion without suppressing liver glucose production.

i) Glucosamine
Systematic name: ***2-Amino-2-deoxy-D-glucose***
Molecular Formula: $\mathbf{C_6H_{13}NO_5}$
Glucosamine is an amino sugar derived from glucose, produced in the body from the sugar glucose and the amino acid glutamine through the action of the enzyme glucosamine synthetase (Fig. 34).

It’s used as a precursor in the biochemical synthesis of glycosylated proteins and lipids. Glucosamine stimulates the synthesis of proteoglycans, glycosaminoglycans (more commonly referred to as mucopolysaccharides), and collagen. It therefore plays a role in the formation of cartilage and the cushioning synovial fluid between the joints; hence it’s

classified as "chondroprotective" agent. Supplementary glucosamine can be an important source of this vital amino sugar for those with reduced capacity to produce glucosamine, such as the elderly.

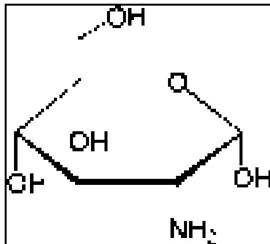


Fig. 34: Glucosamine

Glucosamine acts as an immune modulator with antitumor and antiviral properties, as well as it has some activity against HIV. Oral glucosamine is commonly used for the treatment of osteoarthritis. Since glucosamine is a precursor for glycosaminoglycans, and glycosaminoglycans are a major component of joint cartilage, supplemental glucosamine may help to rebuild cartilage and treat arthritis. Deficiencies or malfunctions in the ability to metabolize this sugar have been linked to diseases of the bowel and bladder. Glucosamine has been shown to help repair the mucosal-lining defensive barrier called the glycosaminoglycan layer (GAG). Defects in the GAG layer have been implimented in Crohn's disease, ulcerative colitis, and interstitial cystitis.

One of the most striking effects of glucosamine is its ability to reduce the progression of experimental cancers. Reductions in blood levels of Glucosamine have been found in those with colon cancer. Distribution of the sugar is also altered when other cancerous tissues are present. It has been found that, with some of the other essential sugars, glucosamine is also vital to learning.

Retinal tissue from human eye donors showed that glucosamine preferred the photoreceptor layer of the retina, suggesting that this sugar is needed in transplantation.

ii) N-Acetylglucosamine (N-Acetyl-D-Glucosamine, or GlcNAc, or NAG)
Systematic name: ***2-(Acetylamino)-2-deoxy-D-glucose***
Molecular formula: $\mathbf{C_8H_{15}NO_6}$
It is a monosaccharide derivative of glucose. Chemically it is an amide between glucosamine and acetic acid; and is significant in several biological systems (Fig. 35).

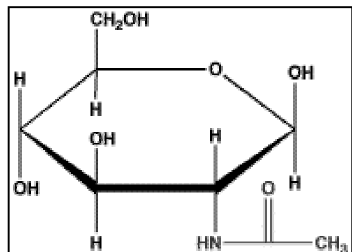


Fig. 35: N-Acetylglucasamine

N-acetylglucosamine is part of a biopolymer in the bacterial cell wall, built from alternating units of GlcNAc and N-acetylmuramic acid (MurNAc), cross-linked with oligopeptides at the lactic acid residue of MurNAc. This layered structure is called peptidoglycan. GlcNAc is the monomeric unit of the polymer chitin, which forms the outer coverings of insects and

crustaceans. GlcNAc is also of note in neurotransmission, where it is thought to be an atypical neurotransmitter functioning in nocioceptive (pain) pathways.

N-acetylglucosamine (GlcNAc) carries out important roles in a broad range of cells from bacteria to humans. One aspect of GlcNAc function is to mediate cellular signaling. In bacteria, GlcNAc induces components that are important for colonization of human hosts, including fimbrins that mediate adhesion to host cells, multidrug exporter genes and Curli fibers that promote biofilm formation. In mammals, GlcNAc is a key sensor of nutrient status that is involved in insulin signaling, cell cycle control, and other essential processes. Nutritional effects that lead to increased GlcNAc levels result in its conversion to UDP-GlcNAc and subsequent attachment to proteins on serine or threonine residues in a dynamic manner that is analogous to modification of proteins by phosphorylation. In fact, the interplay between O-GlcNAc modification and phosphorylation of proteins regulates many critical transcription factors, such as c-myc and p53. O-GlcNAc modification also regulates other processes including proteosome function. In addition to these regulatory roles, GlcNAc contributes to the N-linked glycosylation, glycophosphatidylinositol (GPI) anchor addition to proteins and is polymerized into chitin, which forms part of the fungal cell wall and the exoskeletons of parasites, insects, and other organisms.

N-acetyl-D-glucosamine had been shown to possess an enhancing effect on the production of IgG and IgM in mice. It also plays an important role in nutrient sensing and cellular regulation in a wide range of organisms from bacteria to humans. N-acetylglucosamine receptors are found in the thyroid gland, which indicates that it plays a role in the transport of thyroglobulin (an iodine-containing glycoprotein). Concentrated amounts of N-acetylglucosamine are found in the testes, liver, small intestines, epithelial cells of the endocrine and sebaceous glands, and endothelial cells of blood vessels.

N-AcetylGlucosamine (or glucosamine its metabolic derivative) helps in immune system functioning particular in regards to HIV and tumors. In addition, GlcNAc decreases pain and inflammation and increases range of motion in osteoarthritis patients and helps repair cartilage. GlcNAc has also been implicated as an aid to learning during certain mice studies. The saccharide or its derivative is found in the brains of mammal implying a relation to nerve functioning for learning.

In regards to disease processes, N-acetylglucosamine has been linked to Crohn's disease, interstitial cystitis and ulcerative colitis. Deficiencies of GlcNAc have been linked to diseases of the bowels and bladder. Those with colon cancer show particular deficiencies.

N-acetylglucosamine has been shown to repair the mucosal-lining defensive barrier called the glycosaminoglycan layer (GAG). According to animal studies, N-acetylglucosamine helped prevent the flu virus and herpes virus from occurring. N-acetylglucosamine also has a hand in limiting chloresterol absorption and decreasing insulin secretion. It has been proposed as a treatment for autoimmune diseases.

iii) Galactosamine
Systematic name: ***2-Amino-2-deoxy-D-galactose***
Molecular formula: $\mathbf{C_6H_{13}NO_5}$
Galactosamine is a hexosamine derived from galactose (Fig. 36).

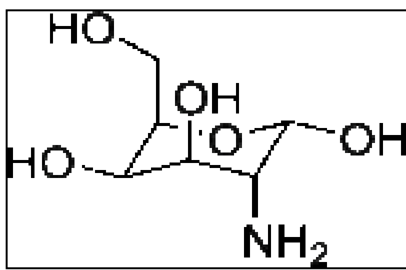


Fig. 36: Galactosamine

This amino sugar is a constituent of some glycoprotein hormones such as follicle-stimulating hormone (FSH) and luteinizing hormone (LH).

iv) N-Acetylgalactosamine (GalNAc, 2-Acetamido-2-deoxy-D-galactopyranose or N-Acetyl-D-galactosamine)
Systematic name: ***2-(Acetylamino)-2-deoxy-D-galactose***
Molecular formula: $\mathbf{C_8H_{15}NO_6}$
It is a monoacetylaminosaccharide derivative of galactose. Chemically it is an amide between galactosamine and acetic acid (Fig. 37).

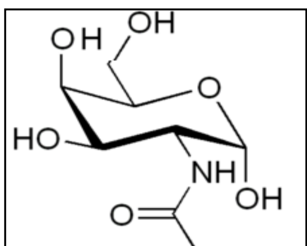


Fig. 37: N-Acetylgalactosamine
In humans it is the terminal carbohydrate forming the antigen of blood group A.

N-Acetylgalactosamine is necessary for intercellular communication, and is concentrated in sensory nerve structures of both humans and animals. This saccharide may inhibit the growth of some tumors.

Lower than normal levels of this sugar have been found in patients with heart disease implying that these conditions may be reversed if a supplementation of N-acetylgalactosamine were to be added to the diet. It appears that this sugar plays a role in joint function, scavenging harmful radicals like superoxide generated during the course of action of macrophage in inflammatory response

N-acetylgalactosamine also seems to play an important role in the immune system. Contained in macrophages and neutrophils, it may play a significant role in the etiolology of joint inflammation and could be important in such conditions as rheumatoid arthritis. N-acetylgalactosamine is localized in the golgi apparatus and ER – found in cell organelles and associated with synthesis of various proteins and enzymes.

It is also found on the surface of cortical neurons and involved in synaptic function of the central nervous system and peripheral brain, suggesting its importance in nerve function. N-acetylgalactosamine is also concentrated in other sensory nerve structures especially in the retina, photoreceptors, optic nerve, and the epithelial pigment of the eyes of both humans and animals. This suggests that it may be extremely important for optimal vision. N-

acetylgalactosamine is distributed to several other tissues, suggesting that it is important in the functional role of these tissues. Some of these tissues include the ducts of the kidney, the testes, the skin, and a variety of other structures including sweat glands, some blood vessel cells, and hair follicles. It is known that concentrations of N-acetylgalactosamine decrease with age.

v) Neuraminic acid & Sialic acid

Systematic name: **(Neuraminic acid)** ***5-amino-3,5-dideoxy- D-glycero-D-galacto-non-2-ulosonic acid***

Molecular formula: $\mathbf{C_9H_{17}N_1O_8}$

Neuraminic acid is a 9-carbon monosaccharide. It may be theoretically visualized as the aldol-condensation product of pyruvic acid and D-mannosamine (2-amino-2-deoxy-mannose). The N- or O-substituted derivatives of neuraminic acid are collectively and commonly known as sialic acids, the predominant one being N-acetylneuraminic acid. The amino group bears either an acetyl or a glycolyl group. The hydroxyl substituents may vary considerably: acetyl, lactyl, methyl, sulfate and phosphate groups have been found (Fig.38).

Fig. 38: Neuraminic Acid

Neuraminic acid does not occur naturally, but many of its derivatives are found widely distributed in animal tissues and in bacteria, especially in glycoproteins and gangliosides. N-acetylneuraminic acid (is widely distributed throughout the tissues of the body (brain, adrenal glands, and the heart). It is found mainly in the glycoproteins and glycolipids and also in many fluids including saliva, urine, cerebrospinal fluid, amniotic fluid, and breast milk.

It is important for brain development, learning, memory and cognitive performance. Cancer cells that can metastasize often have a lot of sialic acid rich glycoproteins. This helps these late stage cancer cells enter the blood stream.

During pregnancy, N-acetylneuraminic acid levels are raised suggesting its importance in the immune system along with other physical and mental development systems for infants. Disrupted N-acetylneuraminic acid metabolisms are seen in infants who are developmentally delayed, show a coarsening of facial features, have enlarged livers and/or spleens and fail to produce skin and hair pigmentation.

Like the other glyconutrients, N-acetylneuraminic acid is important for cellular communication and is an immune system modulator. As an immune modulator, N-acetylneuraminic acid affects the viscosity of mucus, which in turn repels viruses, bacteria and other pathogens. In fact, N-acetylneuraminic acid has been shown to be effective in defending against viruses that cause hepatitis, viral pneumonia and cold sores as well as the common cold. This in turn decreased the severity of asthmatic bronchial spasms and allergic reactions as well. N-acetylneuraminic acid is rapidly expelled from the kidneys and bladder. N-acetylneuraminic acid also lowers the LDL (bad cholesterol) levels and influences blood coagulation.

Sialic acid-rich oligosaccharides on the glycoconjugates found on surface membranes help keep water at the surface of cells. The sialic acid-rich regions contribute to creating a negative charge on the cells surface. Since water is a polar molecule, it has a partial positive charge on both hydrogen molecules, it is attracted to cell surfaces and membranes. This also contributes to cellular fluid uptake. Subjects with Sjogren's syndrome and alcohol dependency show markedly low levels of N-acetylneuraminic acid.

vi) N-Acetylmuramic acid, or MurNAc
Systematic name: ***(R)-2-(acetylamino)-3-O-(1-carboxyethyl)-2-deoxy-D-glucose***
Molecular formula: $\mathbf{C_{11}H_{19}NO_8}$
It is the ether of lactic acid and N-acetylglucosamine (Fig. 39).

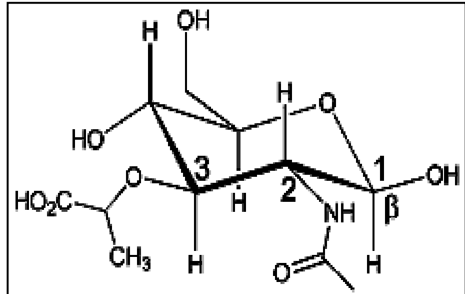


Fig. 39: N-Acetylmuramic Acid

It is part of a biopolymer in the bacterial cell wall, built from alternating units of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) linked by β(1→4) glycosidic bonds, cross-linked with oligopeptides at the lactic acid residue of MurNAc. This layered structure is called peptidoglycan. (see below for specific linkages).

vii) Peptidoglycan
Peptidoglycan, also known as murein, is a polymer consisting of sugars and amino acids that forms a mesh-like layer outside the plasma membrane of eubacteria (Fig. 40a). Structurally, it consists of linear chains of two alternating amino sugars, *N*-acetylglucosamine (GlcNAc or NAG) and *N*-acetylmuramic acid (MurNAc or NAM). The alternating sugars are connected by a β (1, 4) - glycosidic bond.

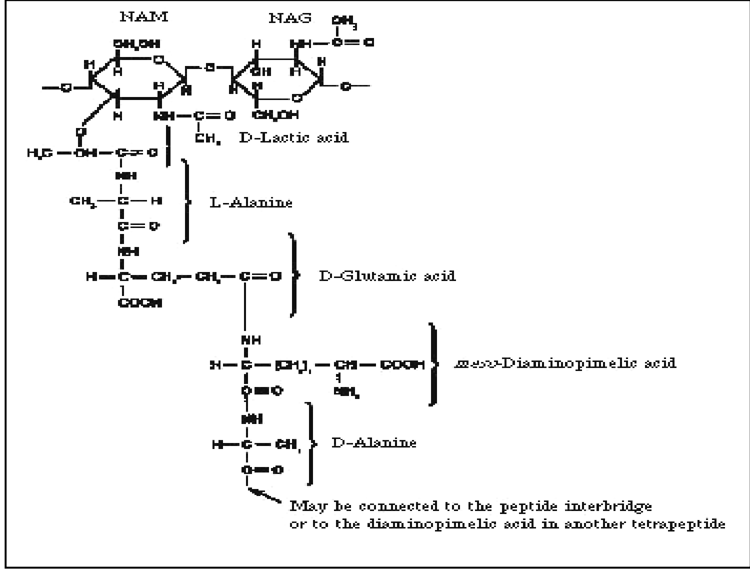


Fig. 40a: Structure of Peptidoglycan layer

Each MurNAc is attached to a short (4- to 5-residue) amino acid chain via its 3rd carbon, normally containing D-alanine, D-glutamic acid, and mesodiaminopimelic acid (Fig. 40b). These three amino acids do not occur in proteins and help protect against attacks by most peptidases. Cross-linking between amino acids in different linear amino sugar chains by an enzyme known as transpeptidase result in a 3-dimensional structure that is strong and rigid. The specific amino acid sequence and molecular structure vary with the bacterial species The peptide chain can be cross-linked to the peptide chain of another strand forming the 3D mesh-like layer. Some Archaea have a similar layer of pseudopeptidoglycan.

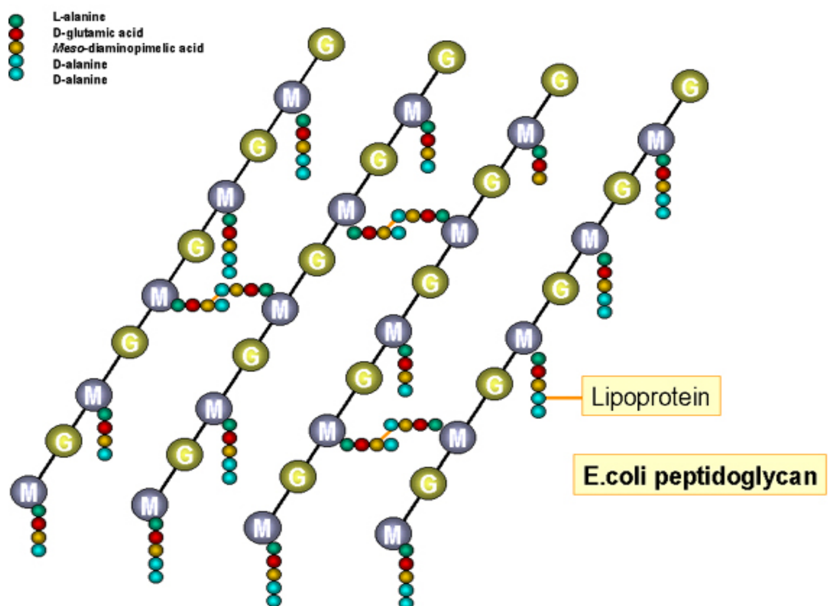


Fig. 40b: Schematic representation of peptidoglycan

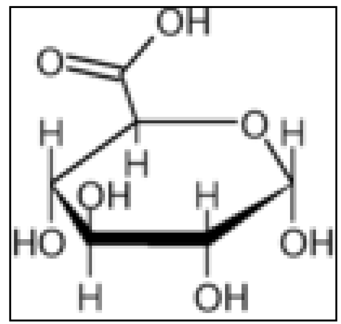
Peptidoglycan serves a structural role in the bacterial cell wall, giving structural strength, as well as counteracting the osmotic pressure of the cytoplasm. A common misconception is that peptidoglycan gives the cell its shape; however, whereas peptidoglycan helps maintain the structure of the cell. Peptidoglycan is also involved in binary fission during bacterial cell reproduction.The peptidoglycan layer is substantially thicker in Gram-positive bacteria (20 to 80 nm) than in Gram-negative bacteria (7 to 8 nm), with the attachment of the S-layer. Peptidoglycan forms around 90% of the dry weight of Gram-positive bacteria but only 10% of Gram-negative strains. In Gram-positive strains, it is important in attachment roles and sterotyping purposes. For both Gram-positive and Gram-negative bacteria, particles of approximately 2 nm can pass through the peptidoglycan.

B. Carboxylic acid sugars

The carbonyl group at C1 of monosaccharides may be oxidized to generate aldonic acids while C-6 oxidation results in uronic acids.

i) Glucuronic acid

Systemic name: ***D-Glucuronic Acid***
Molecular formula: $\mathbf{C_6H_{10}O_7}$
Glucuronic acid is a carboxylic acid. Its structure is similar to glucose. However glucuronic acid's sixth carbon is oxidized to a carboxylic acid. The salts of glucuronic acid are known as glucuronates; the anion $C_6H_9O_7^-$ is the glucuronate ion (Fig. 41).



ig. 41: Glucuronic Acid

Glucuronic acid is highly soluble in water. In the animal body, glucuronic acid is often linked to poisonous substances, mainly in the liver, to allow for subsequent elimination, and to hormones to allow for easier transport. These linkages involve O-glycosidic bonds. The process is known as glucuronidation, and the resulting substances are known as glucuronides (or glucuronosides).

ii) Gluconic acid

Systemic name***: D-Gluconic Acid***

Molecular formula: $\mathbf{C_6H_{12}O_7}$

Gluconic acid is the carboxylic acid formed by the oxidation of the first carbon of glucose and. When dissolved in water, it forms the gluconate ion $C_6H_{11}O_7$; the salts of gluconic acid are also known as gluconates. The chemical structure of gluconic acid consists of a six-carbon chain with five hydroxyl groups terminating in a carboxyl group (Fig. 42). This latter group can lose a hydrogen ion and thus turns the molecule into an acid. In aqueous solution, some gluconic acid molecules will convert to the cyclic ester Glucono delta lactone (Fig. 43), and the two exist in equilibrium

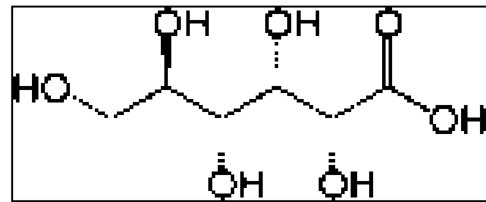


Fig. 42: Gluconic Acid

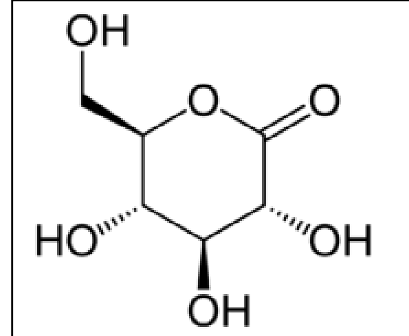


Fig. 43: Glucurono delta lactone

Occurrence & Uses: Gluconic acid occurs naturally in fruit, honey, kombucha tea and wine and is used as a food additive, an acidity regulator. It is a strong chelating agent, especially in alkaline solution.

C. Sugar phosphates

Phosphorylated sugars are another important class of derivatives of sugars. At the normal pH of a cell, most of the hydroxyl groups (OH) on the phosphates are ionized (O-). The hydroxyl group of sugars form ester bond with phosphates.

Phosphorylated sugars are key intermediates in energy generation and biosynthesis. Sugar phosphates are intermediates in monosaccharide metabolism and are building blocks for energy-providing nucleotides and for nucleic acids. Some important sugar phosphates are listed below.

i) Glucose 6 phosphate

Molecular formula: $\mathbf{C_6H_{13}O_9P}$

Glucose 6-phosphate (Fig. 44) (also known as Robison ester) is glucose sugar phosphorylated on carbon 6. This compound is very common in cells as the vast majority of glucose entering a cell is phosphorylated in this way. Because of its prominent position in cellular chemistry, glucose 6-phosphate has many roles within the cell. It lies at the start of two major metabolic pathways: Glycolysis and Pentose phosphate pathway. In addition, it may also be converted to glycogen or starch for storage. This storage is in the liver and muscles in the form of glycogen for most multicellular animals, and in intracellular starch or glycogen granules for most other organisms.

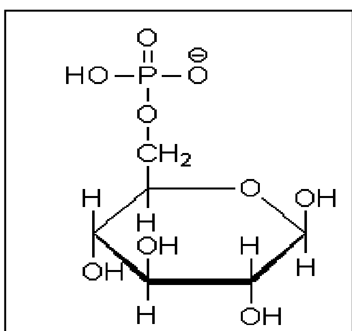


Fig. 44:Glucose 6 phosphate

ii) Adenosine triphosphate (ATP)

Molecular formula: $\mathbf{C_{10}H_{16}N_5O_{13}P_3}$

Adenosine 5'-triphosphate (ATP) (Fig. 45) is a multifunctional nucleotide that is most important as a "molecular currency" of intracellular energy transfer. In this role, ATP transports chemical energy within cells for metabolism. It is produced as an energy source during the processes of photosynthesis and cellular respiration and consumed by many enzymes and a multitude of cellular processes including biosynthetic reactions, motility and cell division. In signal transduction pathways, ATP is used by kinases that phosphorylate proteins and lipids, as well as by adenylate cyclase, which uses ATP to produce the second messenger molecule cyclic AMP.

Fig. 45: Adenosine tri-phosphate

iii) Cyclic Adenosine MonoPhosphate (cAMP)

Molecular formula: $\mathbf{C_{10}H_{12}N_5O_6P}$

Cyclic adenosine monophosphate (cAMP, cyclic AMP or 3'-5'-cyclic adenosine monophosphate) is a molecule that is important in many biological processes (Fig. 46); it is derived from adenosine triphosphate (ATP). cAMP is synthesised from ATP by adenylyl cyclase which is located in the cell membranes. Adenylyl cyclase is activated by the hormones glucagon and epinephrine. cAMP is a second messenger, used for intracellular signal transduction, such as transferring the effects of hormones like glucagon and adrenaline, which cannot get through the cell membrane. Its purposes include the activation of protein kinases and regulating the effects of adrenaline and glucagon. It is also used to regulate the passage of Ca^{2+} through ion channels.

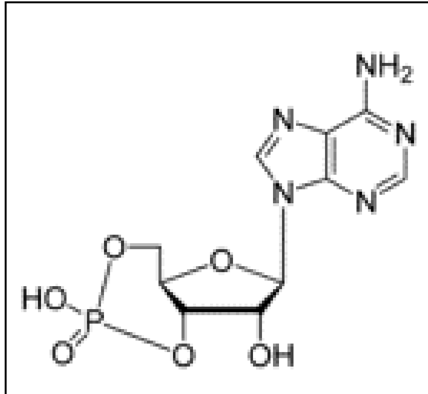


Fig 46:Cyclic Adenosine monophosphate

D. Sugar alcohols

A sugar alcohol (also known as a polyol, polyhydric alcohol, or polyalcohol) is a hydrogenated form of carbohydrate, whose carbonyl group (aldehyde or ketone, reducing sugar) has been reduced to a primary or secondary hydroxyl group. Its general formula is H(HCHO)n+1H, whereas sugar's is H(HCHO)nHCO.Some important sugar alcohols are as follows. Lactitol, sorbitol, xylitol, mannitol, and maltitol are all sugar alcohols. The United States Food and Drug Administration (FDA) classify sugar alcohols as "generally recognized as safe" (GRAS). They are approved as food additives, and are recognized as not

contributing to tooth decay or causing increases in blood glucose. Few sugar alcohols are discussed and illustrated below in table. (Table sugar alcohol)

i) Ethylene glycol
Systemic name: ***1,2-Ethanediol***
Molecular formula: $\mathbf{C_2H_6O_2}$
It is an alcohol with two -OH groups (a diol), a chemical compound.
$\mathbf{C_2H_4O + H_2O = HOCH_2CH_2OH}$
In its pure form, it is an odorless, colorless, syrupy liquid with a sweet taste. Ethylene glycol is toxic, and its ingestion should be considered a medical emergency. (a Table sugar alcohol)

ii) Glycerol
Systemic name: ***1,2,3-Propanetriol***
Molecular formula: $\mathbf{C_3H_8O_3}$
Glycerol has three hydrophilic alcoholic hydroxyl groups that are responsible for its solubility in water and its hygroscopic nature. This colorless, odorless, viscous liquid is widely used in pharmaceutical formulations. Also commonly called glycerin or glycerine, it is a sugar alcohol, and is sweet-tasting and of low toxicity. It is a central component of lipids. Glycerol is a precursor for synthesis of triacylglycerols and of phospholipids in the liver and adipose tissue. When the body uses stored fat as a source of energy, glycerol and fatty acids are released into the bloodstream. The glycerol component can be converted to glucose by the liver and provides energy for cellular metabolism. (b Table sugar alcohol)

Table: Sugar alcohols

Sugar Alcohols	Structure
a. Ethylene glycol	
b. Glycerol	
c. Erythritol	
d. Xylitol	
e. Mannitol	

f. Sorbitol	
g. Maltitol	
h. Lactitol	

iii) Erythritol
Systemic name: ***(2R, 3S)-butane-1,2,3,4-tetraol***
Molecular formula**:** $\mathbf{C_4H_{10}O_4}$
Erythritol is alcohol derivative of erythrose. Erythritol is a natural sugar alcohol (a type of sugar substitute), which has been approved for use in the United States and throughout much of the world. It occurs naturally in fruits and fermented foods. At industrial level, it is produced from glucose by fermentation with yeast, *Moniliella pollinis*. (c Table sugar alcohol)

iv) Xylitol
Systemic name: ***Xylo-Pentane-1, 2, 3, 4, 5-pentol***
Molecular formula: $\mathbf{C_5H_{12}O_5}$
Xylitol is derived from xylose. It is a five-carbon sugar alcohol that is used as a sugar substitute. It naturally occurs in the fibers of many fruits and vegetables, including various berries, cornhusks, oats, and mushrooms. Xylitol is roughly as sweet as sucrose but with only two-thirds the food energy.

Possessing approximately 40% less food energy, xylitol is a low-calorie alternative to table sugar. It is a "toothfriendly" sugar and may actively aid in repairing minor cavities caused by dental caries. Xylitol also appears to have potential as a treatment for osteoporosis. The open nature of xylitol and its ability to form many different sugar-like structures appears to interfere with the ability of many bacteria to adhere. (d Table sugar alcohol)

v) Mannitol
Systemic name: ***Xylo-Pentane-1, 2, 3, 4, 5-pentol***
Molecular formula: $\mathbf{C_5H_{14}O_6}$
Mannitol is a sorbitol stereoisomer. It is an osmotic diuretic agent and a weak renal vasodilator. It was originally isolated from the secretions of the Flowering Ash, called Manna after their resemblance to the Biblical food, and may also be referred to as Mannite and Manna Sugar. (e Table sugar alcohol)

Mannitol is used clinically to reduce acutely raised intracranial pressure, until more definitive treatment can be given, and to treat patients with oliguric renal failure. Mannitol can also be used to open the blood-brain barrier by temporarily shrinking the tightly coupled endothelial cells that make up the barrier. This makes mannitol indispensable for delivering various drugs directly to the brain (e.g. in the treatment of Alzheimer's disease). Mannitol is also used as a sweetener for people with diabetes. Since mannitol has a negative heat of solution, it is used as a sweetener in "breath-freshening" candies. It is sometimes used as an adulterant or cutting agent for heroin, methamphetamines or other illicit drugs. It is a non-permeating molecule; i.e., it cannot cross biological membranes. Mannitol is commonly used in the circuit prime of a heart lung machine during cardiopulmonary bypass (CPB). The presence of mannitol preserves renal function during the times of low blood flow and pressure, while the patient is on bypass. The solution prevents the swelling of endothelial cells in the kidney, which may have otherwise reduced blood flow to this area and resulted in cell damage. Mannitol is also being developed as a treatment for cystic fibrosis and bronchiectasis and as a diagnostic test for airway hyperresponsiveness. The mannitol is orally inhaled as a dry powder through what is known as an osmohaler.

vi) Sorbitol
Systemic name: ***D-Glucitol***
Molecular formula: $\mathbf{C_6H_{14}O_6}$
Sorbitol, also known as glucitol, is a sugar alcohol (Table f) metabolized slowly by the body. It is obtained by reduction of glucose changing the aldehyde group to an additional hydroxyl group hence the name sugar alcohol. (f Table sugar alcohol)

Sorbitol is a sugar substitute often used in diet foods and sugar-free chewing gum. It occurs naturally in many stone fruits and berries from trees of the genus Sorbus. Sorbitol is a nutritive sweetener and provides 65% dietary energy, while retaining 60% of the sweetness. As a food additive it is categorized as a sweetener, emulsifier and humectant. Sorbitol can be used as a non-stimulant laxative as either an oral suspension or suppository. Ingesting large amounts of sorbitol can lead to some abdominal pain, gas, and mild to severe diarrhea. Even in the absence of dietary sorbitol, cells also produce sorbitol naturally. Diabetic retinopathy and neuropathy may be related to excess sorbitol in the cells of the eyes and nerves. The source of this sorbitol in diabetics is excess glucose, which goes through the polyol pathway. Sorbitol is often used in modern cosmetics as a humectant and thickener. Some transparent gels can only be made with sorbitol as it has a refractive index sufficiently high for transparent formulations. It is also used as a cryoprotectant additive (mixed with sucrose and sodium polyphosphates) in the manufacture of surimi, a highly refined, uncooked fish paste most commonly produced from Alaska (or walleye) pollock (Theragra chalcogramma). Sorbitol is identified as a potential key chemical intermediate from biomass resources. Complete reduction of sorbitol opens the way to alkanes such as hexane, which can be used as a biofuel.

vii) Maltitol
Systemic name: ***4-O-α-D-Glucopyranosyl-D-glucitol***
Molecular formula: $\mathbf{C_{12}H_{24}O_{11}}$
Maltitol is a sugar alcohol (a polyol) used as a sugar substitute (g Table sugar alcohol) because it has fewer calories, does not promote tooth decay and has a somewhat lesser effect on blood glucose. It has 90% of the sweetness of sucrose (table sugar) and nearly identical properties, except for browning. Unfortunately, maltitol is well known to cause gastric distress if consumed in great quantities. Chemically, maltitol is also known as 4-O-a-

glucopyranosyl-D-sorbitol. Commercially, maltitol is a disaccharide produced by Corn Products It is made by hydrogenation of maltose obtained from starch.

It is used especially in production of sweets: sugarless hard candies, chewing gum, chocolates, baked goods, and ice cream. It is not metabolized by oral bacteria, so it does not promote tooth decay. It is somewhat more slowly absorbed than sucrose, which makes it somewhat more suitable for people with diabetes than sucrose. Due to its slow absorption, excessive consumption can have laxative effect and sometimes can cause gas and/or bloating. This means that maltitol is particularly associated with gastric issues.

viii) Lactitol
Systemic name: ***4-O- β-D-galactocopyranosyl-D-glucitol***
Molecular formula: $C_{12}H_{24}O_{11}$
Lactitol is a sugar alcohol used as a replacement bulk sweetener for low calorie foods with approximately 40% of the sweetness of sugar. (h Table sugar alcohol)

Lactitol is used in a variety of low food energy or low fat foods. High stability makes it popular for baking. It is used in sugar-free candies, cookies (biscuits), chocolate, and ice cream. Lactitol also promotes colon health as a prebiotic. Lactitol only has 60% calories, compared for typical carbohydrates. Lactitol is also approved for use in foods in most countries around the world, though lactitol can cause cramping, flatulence, and diarrhoea in some individuals.

2. Derivatives of polysaccharides

(Glycosaminoglycans (GAGs) or mucopolysaccharides)

Classification: Glycosaminoglycans (GAGs) or mucopolysaccharides are long unbranched anionic polysaccharides consisting of a repeating disaccharide unit (sugar 1 and sugar 2) **Table below**. Members of the glycosaminoglycan family vary in the type of hexosamine, hexose or hexuronic acid unit they contain (e.g. glucuronic acid, iduronic acid, galactose, galactosamine, glucosamine). They also vary in the geometry of the glycosidic linkage. They are synthesized in either endoplasmic reticulum and/or Golgi.

Examples of GAGs:

Name	Sugar 1	Sugar 2	Linkage	Unique features
Chondroitin sulphate	N-acetylgalactosamine	Glucuronic acid	β (1→3)	Most prevalent GAG
Dermatan sulphate	Iduronic acid	N_Acetylgalactosamine	β (1→3)	Only one with iduronic acid
Keratan sulphate	Galactose	(varies)	β (1→4)	Very variable
Heparin	Glucuronic acid	Glucosamine	α (1→4)	Only one intracellular; high negative charge density
Heparan sulphate	Glucuronic acid	Glucosamine	α (1→4)	Similar to heparin but extracellular
Hyaluronan	Glucuronic acid	N-Acetylglucosamine	β (1→3)	Only bacterial one, only one without sulfur

A glycosaminoglycans unit consists of an N-acetyl-hexosamine and a hexose or hexuronic acid, either or both of which may be sulfated. The combination of the sulfate group and the carboxylate groups of the uronic acid residues gives them a very high density of negative charge.

This family of carbohydrates is essential or important for the life of vertebrates and an assortment of lower animals. It is present on the cell surface and in the extracellular matrix of animals. GAGs form an important component of connective tissues. GAG chains may be covalently linked to a protein to form proteoglycans.

i) Chondroitin sulfate
Chondroitin sulfate is a sulfated glycosaminoglycan (GAG) composed of a chain of alternating sugars (N-acetylgalactosamine and glucuronic acid) (Fig. 47). It is usually found attached to proteins as part of a proteoglycan. A chondroitin chain can have over 100 individual sugars, each of which can be sulfated in variable positions and quantities. Understanding the functions of such diversity in chondroitin sulfate and related glycosaminoglycans is a major goal of glycobiology. Chondroitin sulfate is an important structural component of cartilage and provides much of its resistance to compression. Chondroitin sulfate chains are unbranched polysaccharides of variable length containing two alternating monosaccharides: D-glucuronic acid (GlcA) and N-acetyl-D-galactosamine (GalNac). Some GlcA residues are epimerized into L-iduronic acid (IdoA); the resulting disaccharide is then referred to as dermatan sulfate.

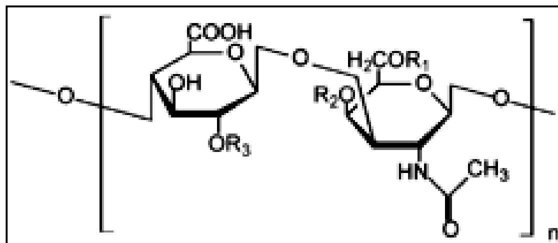


Fig. 47: Chondroitin sulfate

Protein attachment: Chondroitin sulfate chains are linked to hydroxyl groups on serine residues of certain proteins. Exactly how proteins are selected for attachment of glycosaminoglycans is not understood. Empirically, glycosylated serines are often followed by a glycine and have neighboring acidic residues, but this motif does not always predict glycosylation.

Attachment of the GAG chain begins with four monosaccharides in a fixed pattern: Xyl - Gal - Gal - GlcA. Each sugar is attached by a specific enzyme, allowing for multiple levels of control over GAG synthesis. Xylose begins to be attached to proteins in the endoplasmic reticulum, while the rest of the sugars are attached in the Golgi apparatus.

Sulfation: Each monosaccharide may be left unsulfated, sulfated once, or sulfated twice. Most commonly the hydroxyls of the 4 and 6 positions of the N-acetyl-galactosamine are sulfated. Sulfation is mediated by specific sulfotransferases.

Functions: Chondroitin's functions largely depend on the properties of the overall proteoglycan of which it is a part. These functions can be broadly divided into structural and

regulatory roles. However, this division is not absolute and some proteoglycans have both structural and regulatory roles.
Structural: Chondroitin sulfate is a major component of extracellular matrix, and is important in maintaining the structural integrity of the tissue. This function is typical of the large aggregating proteoglycans: aggrecan, versican, brevican, and neurocan. As part of aggrecan, chondroitin sulfate is a major component of cartilage. The tightly packed and highly charged sulfate groups of chondroitin sulfate generate electrostatic repulsion that provides much of the resistance of cartilage to compression. Loss of chondroitin sulfate from the cartilage is a major cause of osteoarthritis.

Regulatory: Chondroitin sulfate readily interacts with proteins in the extracellular matrix due to its negative charges. These interactions are important for regulating a diverse array of cellular activities. In the nervous system, chondroitin sulfate proteoglycans regulate the growth and development of the nervous system as well as the nervous system response to injury.

Medical use: Chondroitin is an ingredient found commonly in dietary supplements used as an alternative medicine to treat osteoarthritis. It is commonly sold together with glucosamine.

ii) Dermatan sulfate
Dermatan sulfate is a glycosaminoglycan (formerly called a mucopolysaccharide) found mostly in skin, but also in blood vessels, heart valves, tendons, and lungs.

Function: Dermatan sulfate may have roles in coagulation, cardiovascular disease, carcinogenesis, infection, wound repair, and fibrosis. Dermatan sulfate accumulates abnormally in several of the mucopolysaccharidosis disorders.

iii) Keratan sulfate
Keratan sulfate is any of several sulfated glycosaminoglycans that have been found especially in the cornea, cartilage, and bone. It is a large, highly hydrated structural carbohydrates molecule, which in joints can act as a cushion. There are two main types. Type I is found in the cornea, and Type II is found in cartilage.

iv) Heparin
Heparin is a member of the glycosaminoglycan family of carbohydrates (which includes the closely related molecule heparan sulfate). It is widely used as an injectable anticoagulant and has the highest negative charge density of any known biological molecule. It can also be used to form an inner anticoagulant surface on various experimental and medical devices such as test tubes and renal dialysis machines.

Structure: Native heparin is a polymer with a molecular weight ranging from 3 kDa to 40 kDa although the average molecular weight of most commercial heparin preparations is in the range of 12 kDa to 15 kDa. Heparin is a member of the glycosaminoglycan family of carbohydrates (which includes the closely related molecule heparan sulfate) and consists of a variably sulfated repeating disaccharide unit (Fig. 48). The main disaccharide units that occur in heparin are shown below. The most common disaccharide unit is composed of a 2-O-sulfated iduronic acid and 6-O-sulfated, N-sulfated glucosamine, IdoA(2S)-GlcNS(6S). Under physiological conditions the ester and amide sulfate groups are deprotonated and attract positively charged counterions to form a heparin salt. It is in this form that heparin is usually administered as an anticoagulant.

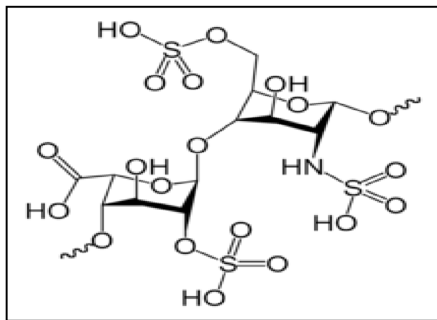


Fig. 48: Heparin

Mechanism of action: Heparin is a naturally occurring anticoagulant produced by basophils and mast cells. Heparin binds to the enzyme inhibitor antithrombin III (AT-III) causing a conformational changes which results in its active site being exposed. The activated AT-III then inactivates thrombin and other proteases involved in blood clotting, most notably factor Xa.

Heparin's exact physiological role is still unclear, because blood anti-coagulation is mostly achieved by endothelial cell-derived heparan sulfate proteoglycans. Heparin is usually stored within the secretory granules of mast cells and only released into the vasculature at sites of tissue injury. It has been proposed that rather than anticoagulation the main purpose of heparin is in a defensive mechanism at sites of tissue injury against invading bacteria and other foreign materials.

v) Heparan sulfate

Heparan sulfate (HS) is a linear polysaccharide found in all animal tissues. It occurs as a proteoglycan (PG) in which two or three HS chains are attached in close proximity to cell suface or extracellular matrix proteins. It is in this form that HS binds to a variety of protein ligands and regulates a wide variety of biological activities, including developmental processes, angiogenesis, blood coagulation and tumour metastasis.

HS structure and differences from heparin: Heparan sulfate is a member of the glycosaminoglycan family of carbohydrates and is very closely related in structure to heparin. Both consist of a variably sulfated repeating disaccharide unit. The most common disaccharide unit within heparan sulfate is composed of a glucuronic (GlcA) linked to N-acetyl glucosamine (GlcNAc) typically making up around 50% of the total disaccharide units. It has been suggested that a GAG should qualify as heparin only if its content of N-sulfate groups largely exceeds that of N-acetyl groups and the concentration of O-sulfate groups exceeds those of N-sulfate (Fig. 49). Under physiological conditions the ester and amide sulfate groups are deprotonated and attract positively charged counterions to form a salt. It is in this form that HS is thought to exist at the cell surface. Heparan sulfate is initially synthesized on a membrane-embedded core protein as a polymer of alternating glucuronate and N-acetylglucosamine residues. Later, in segments of the polymer, glucuronate residues may be converted to the sulfated sugar iduronic acid, while N-acetylglucosamine residues may be deacetylated and/or sulfated. Some cell surface heparan sulfate glycosaminoglycans remain covalently linked to core proteins embedded in the plasma membrane. Proteins involved in signaling and adhesion at the cell surface recognize and bind segments of heparan sulfate chains having particular patterns of sulfation (Fig. 50).

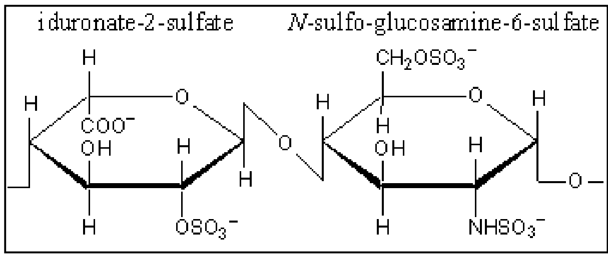


Fig. 49: Heparan sulfate

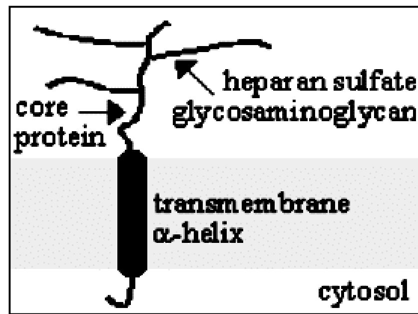


Fig. 50: Heparan sulfate chains

vi) Hyaluronan

Hyaluronan (also called hyaluronic acid or hyaluronate) is a non-sulfated glycosaminoglycan distributed widely throughout connective, epithelial, and neural tissues. Hyaluronan is a polymer of disaccharides themselves composed of D-glucuronic acid and D-N-acetylglucosamine, linked together via alternating β-1,4 and β-1,3 glycosidic bonds. Hyaluronan can be 25,000 disaccharide repeats in length. Polymers of hyaluronan can range in size from 5,000 to 20,000,000 Da *in vivo*. The average molecular weight in human synovial fluid is 3−4,000,000 Da and hyaluronan purified from human umbilical cord is 3,140,000 Da.

It is one of the chief components of the extracellular matrix, contributes significantly to cell proliferation and migration, and may also be involved in the progression of some malignant tumors. The average 70 kg man has roughly 15 grams of hyaluronan in his body, one third of which is turned over (degraded and synthesised) every day.

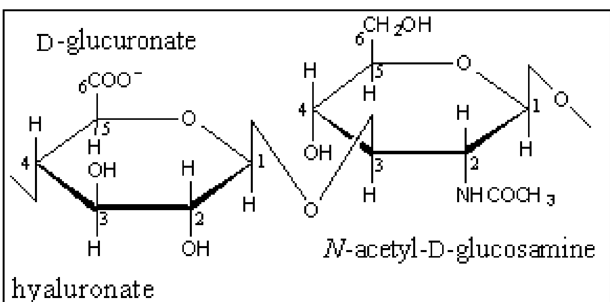


Fig. 51: Hyaluronan

Functions: Until the late 1970s, hyaluronan was described as a "goo" molecule, a ubiquitous carbohydrate polymer that is part of the extracellular matrix. For example, hyaluronan is a major component of the synovial fluid and was found to increase the

viscosity of the fluid. Along with lubricin, it is one of the fluid's main lubricating components. Hyaluronan is an important component of articular cartilage, where it is present as a coat around each cell (chondrocyte). When aggrecan monomers bind to hyaluronan in the presence of link protein, large highly negatively charged aggregates form. These aggregates imbibe water and are responsible for the resilience of cartilage (its resistance to compression). The molecular weight (size) of hyaluronan in cartilage decreases with age however the amount increases.

Hyaluronan is also a major component of skin, where it is involved in tissue repair. When skin is excessively exposed to UV-B rays, it becomes inflamed (sunburn) and the cells in the dermis stop producing as much hyaluronan and increase the rate of its degradation. Hyaluronan degradation products also accumulate in the skin after UV exposure.

While it is abundant in extracellular matrices, hyaluronan also contributes to tissue hydrodynamics, movement and proliferation of cells, and participates in a number of cell surface receptor interactions, notably those including its primary receptor, CD44. CD44 participates in cell adhesion interactions required by tumor cells.

Carbohydrates in living system

A: ABO Blood group system

The International Society of Blood Transfusion currently recognizes 29 blood group systems (including the ABO and Rh systems). ABO blood types are also present in some animals, for example; apes such as chimpanzees, bonobos and gorillas.

Carbohydrates form the most important part of blood group system in human blood transfusion. The associated anti-A antibodies and anti-B antibodies are usually IgM antibodies, which are usually produced in the first years of life by sensitization to environmental substances such as food, bacteria and viruses. The corresponding blood group carbohydrate structures, designated ABH, are found at the termini of oligosaccharide chains on glycoproteins and glycolipids on the surface of RBC & endothelial & most epithelial cells. The immunodominant monosaccharide that determines blood group A, is a terminal α-1,3-linked N-acetylgalactosamine (GalNAc), whereas the corresponding monosaccharide of blood group B specificity is an α-1,3-linked galactose. Group O cells lack chains, which instead are terminated with α-1,2-linked-fucose.

The ABO blood group system is widely credited to have been discovered by the Austrian scientist Karl Landsteiner, who found three different blood types A, B, and O in 1900; Czech serologist Jan Janský independently pioneered the classification of human blood into four groups. Decastrello and Sturli discovered the fourth type, AB, in 1902. Ludwik Hirszfeld and E. von Dungern discovered the heritability of ABO blood groups in 1910-11, with Felix Bernstein demonstrating the correct blood group inheritance pattern of multiple alleles at one locus in 1924.

Chemical nature & Structure: The antigens, which determine blood types belong to glycoproteins and glycolipids. There are three types of blood-group antigens: O, A, and B. They differ only slightly in the composition of carbohydrates. The A antigen and the

B antigen are derived from a common precursor known as the H antigen (or H substance). The H antigen is a carbohydrate sequence with carbohydrates linked mainly to protein (with a minor fraction attached to ceramide moiety). The majority of the ABO determinants are expressed on the ends of long polylactosamine chains attached mainly to Band 3 protein

(1), the anion exchange protein of the red cell membrane, and a minority of the epitopes is expressed on neutral glycosphingolipids (1). In blood group O, the H antigen remains unchanged and consists of a chain of beta-D-galactose, beta-D-N-Acetylglucosamine, beta-D-galactose, and 2-linked, alpha-L-fucose, the chain being attached to the protein or ceramide. H antigens can be changed into A or B antigens by enzymes coded by the blood group A or B genes, which are sugar (glycosyl) tranferases. Type A has an extra alpha-N-Acetyl-D-galactosamine bonded to the D-galactose at the end, while type B has an extra alpha-D-galactose bonded to the D-galactose at the end.

- Individuals with Type A blood can receive blood from donors of type A and type O blood.
- Individuals with type B blood can receive blood from donors of type B and type O blood.
- Individuals with type AB blood can receive blood from donors of type A, type B, type AB, or type O blood. Type AB blood is referred to as the universal recipient.
- Individuals with of O blood can receive blood from donors of only type O.
- Individuals of type A, B, AB and O blood can receive blood from donors of type O blood. Type O blood is called the universal donor.
- Antibodies are not formed against the H antigen except by those with the Bombay phenotype.

In ABH secretors, ABH antigens are secreted by most mucus-producing cells of the body interfacing with the environment, including lung, skin, liver, pancreas, stomach, intestines, ovaries and prostate.

Blood-group antigens

All humans contain enzymes, which catalyze the synthesis of the O antigen. Humans with A-type blood also contain an additional enzyme (called A-type enzyme here), which adds N-acetylgalactosamine to the O antigen. Humans with B-type blood contain another enzyme (called B-type enzyme here), which adds galactose to the O antigen. Humans with AB-type blood contain both A-type and B-type enzymes while humans with O-type blood lack both types of enzymes (Fig. 52).

The Rh System

Rh antigens are transmembrane proteins with loops exposed at the surface of red blood cells. They appear to be used for the transport of carbon dioxide and/or ammonia across the plasma membrane. They are named for the rhesus monkey in which they were first discovered. There are a number of Rh antigens. Red cells that are "Rh positive" are designated as D. About 15% of the population have no RhD antigens and thus are "Rh negative". The major importance of the Rh system for human health is to avoid the danger of RhD incompatibility between mother and fetus.

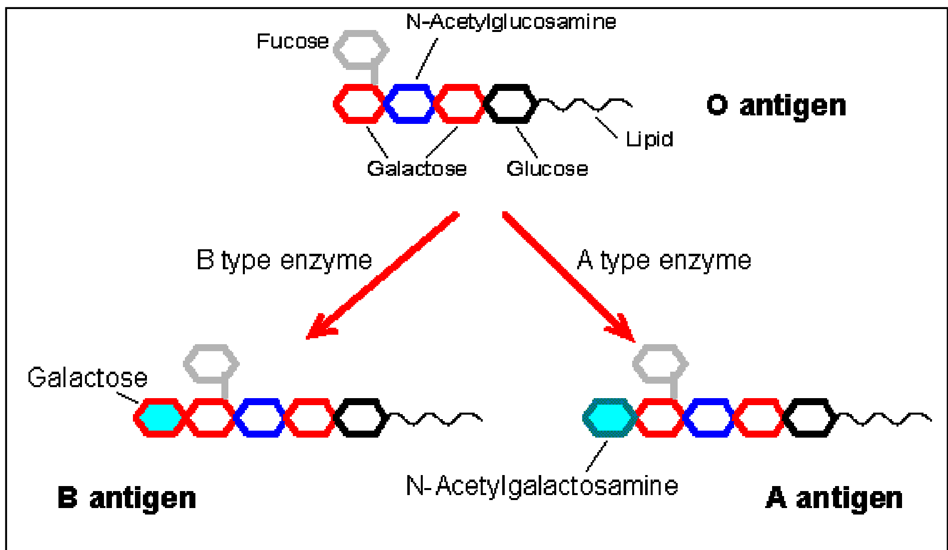


Fig. 52: Blood group antigens

During birth, there is often a leakage of the baby's red blood cells into the mother's circulation. If the baby is Rh positive (having inherited the trait from its father) and the mother Rh-negative, these red cells will cause her to develop antibodies against the Rh D antigen. The antibodies, usually of the IgG class, do not cause any problems for that child, but can cross the placenta and attack the red cells of a subsequent Rh+ fetus. This destroys the red cells producing anemia and jaundice. The disease, called erythroblastosis fetalis or hemolytic disease of the newborn may be so severe as to kill the fetus or even the newborn infant. It is an example of an antibody-mediated cytotoxicity disorder.

B: Sugars in the cell wall of bacteria

In contrast to eukaryotic cells, bacterial cells have a cell wall in addition to a lipid bilayer membrane. These are essentially carbohydrate polymers, which offer protection from exterior hypotonic condition and the high internal osmotic pressures, preventing swelling and bursting of the cells. The membrane consists of a peptidoglycan (Fig. 53).

a) In Gram positive bacteria

Gram-positive bacteria can be stained with Gram stain. The wall consists of a GlcNAc (b 1->4) MurNAc repeat (like that in chitin which is a polymer of GlcNAc in (b 1->4) links, but in which the OH of lactate is in ether-linkage to C3 to form N-Acetylmuramic acid). A tetrapeptide (Ala-D-isoGlu-Lys-D-Ala) is attached in amide link to the carboxyl group of the lactate in MurNAc. The GlcNAc (b 1->4) MurNAc strands are covalently connected by a pentaglycine bridge through the epsilon amino group of the tetrapeptide Lys on one strand and the D-Ala of a tetrapeptide on another strand (Fig. 54).

Techioic acids are often attached to the C6 of MurNAc. Teichoic acid is a polymer of glycerol or ribitol to which alternative GlcNAc and D-Ala are linked to the middle C of the glycerol. Multiple glycerols are linked through phosphodiester bonds. These teichoic acids often make up 50% of the dry weight of the cell wall, and present a foreign (or

antigenic) surface to infected hosts. These often serve as receptors for viruses that infect bacteria (bacteriophages)

.

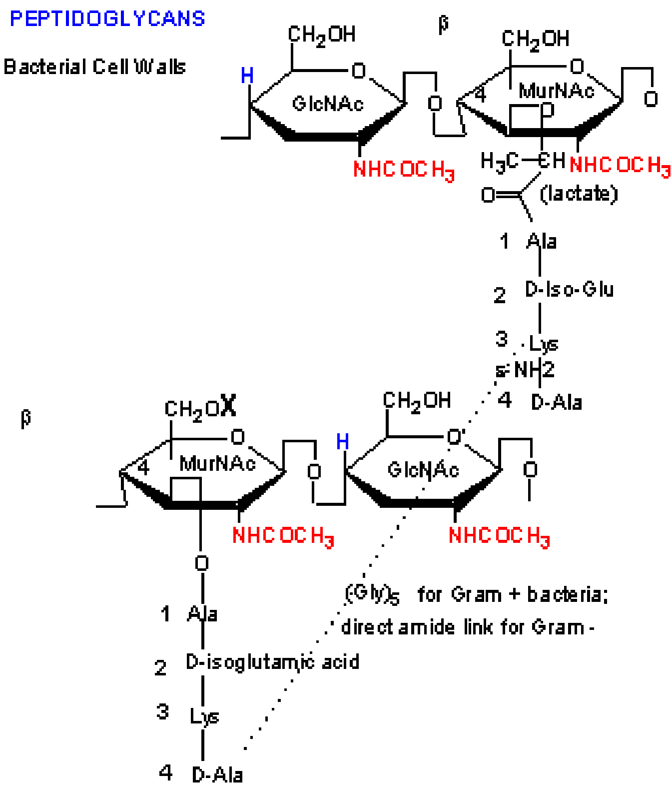


Fig. 53: Bacterial cell wall

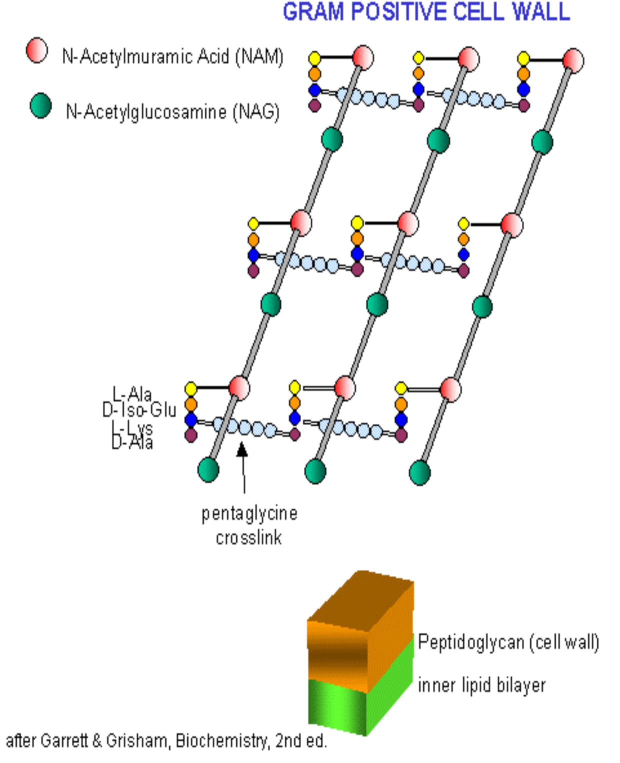


Fig. 54: Gram positive bacterial cell wall

b) In Gram negative bacteria

These bacterial cannot be stained with Gram stain. The wall consists of the same structure as in Gram positive bacteria, but the GlcNAc (b 1->4) MurNAc strands are covalently connected through a direct amide bond between the epsilon amino group of the tetrapeptide Lys on one strand and the D-Ala of a tetrapeptide on another strand. (i.e. no pentaGly spacer). In addition, Gram negative bacteria don't have teichoic acid polymers. Rather they have a second, outer lipid bilayer. The cell wall is sandwiched between the inner and outer bilayers. The space between the lipid bilayers is called the periplasmic space. A hydrophobic protein covalently attaches (through an amide link from a protein Lys) to the cell wall at the last amino acid in the tetrapeptide unit of the wall (actually diaminopimelic acid which replaces about 10% of the D-Ala in the cell wall). The N-terminal of the hydrophobic proteins attaches to the outer lipid membrane through a Ser (Fig. 55). The outer membrane is coated with a lipopolysaccharide (LPS) of varying composition. The LPS determines the antigenicity of the bacteria. The different LPS are called the O-antigens (Fig. 56).

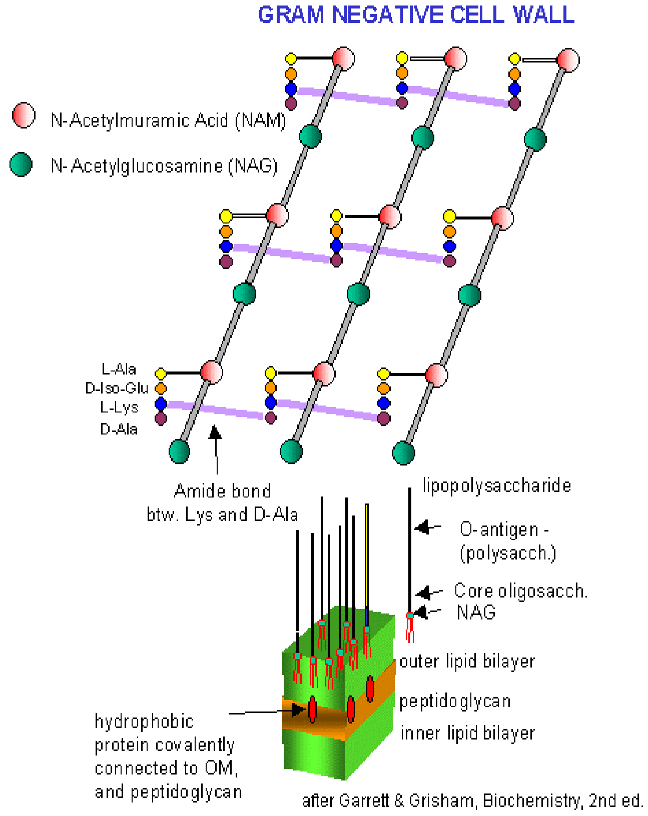


Fig. 55: Gram Negative bacterial cell wall

C: Sugars in the cell wall of fungus

The composition of the fungal cell wall is characterized as a relatively simple structure made up of "cellulose" and chitin. The fungal cell wall can make up 30% or more of the dry weight of the fungus, and the fungi are characterized by external digestion of food followed by selective absorption of the digestion products. So the fungal cell walls are a complex, specialized and highly dynamic system, constantly being regenerated and remodeled according to the needs. Therefore, many of the cell wall-associated proteins

enzymes do hydrolyze chitin and polysaccharides. The fungal cell wall is very different from insect exoskeletons or plant cell walls, which are terminally differentiated structures.

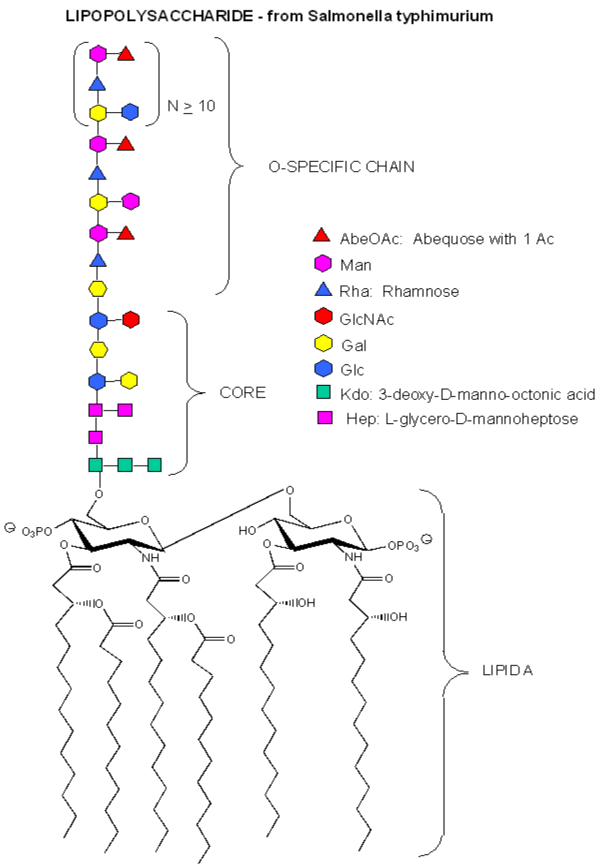


Fig. 56: Liposaccharide organization in bacteria

The cell wall is generally constructed of three layers: (1) An α-glucan layer (a glucan is a polymer of glucose), (2) A β-glucan layer, and (3) An outer layer of glycoprotein.

In addition, chitin may be a significant component of certain cell wall structures. The α-glucan layer, if present, is generally composed of the α (1→3)-glucan polymer. However, α (1→4) glycosides are variably present. Compare glycogen, which is α (1→4)-glucan with (1→6) side chains. Where present, the α-glucan material appears as a fibrillar layer adjacent to the plasma membrane and is thought to serve a largely structural role, stiffening the basal layer of the cell wall. The α-glucan layer is rarely represented in diagrams of the fungal cell wall because it does not occur in *Saccharomyces*, the usual model system. Among ascomycetes, the alpha glucan is found in *Schizosaccharomyces,* but is not known from any other yeasts. The material is common among all groups in the *Pezizomycotina*. However, in *Lecanoromycetes*, a very large proportion tends to be in the α (1→4) form. Alpha glucans form a significant part of the cell wall in many basidiomycetes.

The bulk material of the cell wall is usually in the form of β (1→3)-glucan. This forms a very stable hydrogen-bonded triple helix in solution, and probably *in vivo*. The packing of these triple helix structures is controlled by the size and frequency of very short (1→6) side chains, sometimes consisting of a single glucose monomer.

In addition to β(1→3)-glucan, the cell wall contains β(1→6)-glucan. It may be described as a β(1→3)-glucan with big side-chains, with a true β(1→6) backbone. This material may be peripheral to the bulk β(1→3)-glucan and is strongly involved in cross-linking the various components of the cell wall, as shown in the figure.

The outermost layer of the cell wall is composed of diverse proteins with polysaccharide side chains composed of mannose. These are attached through their mannan side chains via a (1→3) linkage with the β(1→6)-glucan. In reality the structure is much more complex, involving a wide variety of different interactions between glycoproteins and bulk cell wall materials.

Finally, the fungal cell wall contains also contains variable amounts of chitin. In many systems chitin is a major constituent of the cell wall. In others, it is involved only in cell division or reproductive structures or is virtually absent otherwise.

D: Derivatives with proteins

Sugars combine with proteins in different ways to create glycoproteins, proteoglycans etc.

a. Glycoproteins

A glycoprotein is a biomolecule composed of a protein and a carbohydrate (an oligosaccharide). The carbohydrate is attached to the protein in a co-translational or post-translational modification. This process is known as glycosylation. The addition of sugar chains can happen either at asparagines (N-glycosylation); or at hydroxylysine, hydroxyproline, serine, or threonine (O-glycosylation). Monosaccharides commonly found in eukaryotic glycoproteins include glucose, N-acetylglucosamine, galactose, N-acetylgalactosamine, mannose, fucose, xylose and N-acetylneuraminic acid (also known as sialic acid). In proteins that have segments extending extracellularly, the extracellular segments are often glycosylated.

Important facts

a) Glycoproteins have carbohydrate attached to them and this attachment is a covalent linkage to:

- The hydroxyl (-OH) group of the R group of serine or threonine - called "O-linked" in both cases.
- The amino group ($-NH_2$) in the R group of asparagine - called "N-linked".

b) The carbohydrate consists of short, usually branched, chains of

- Simple sugars (e.g., glucose, galactose)
- Amino sugars (sugars with an amino group, e.g., N-acetylglucosamine), and
- Acidic sugars (sugars with a carboxyl group, e.g., sialic acid)

Sugars are very hydrophilic thanks to their many -OH groups. Their presence makes glycoproteins far more hydrophilic than they would be otherwise and essential for the proper folding of the protein into its tertiary structure. Most of the proteins exposed to the watery surroundings at the surface of cells are glycoproteins. Soluble glycoproteins often show a high viscosity, for example, in egg white and blood plasma.

Types of glycoproteins

Broadly, they can be classified into two types according to the site of glycosylation.

A) In **N-glycosylation** (Fig. 57), the addition of sugar chains can happen at the amide nitrogen on the side chain of the asparagines.

N-linked glycosylation is important for the folding of some eukaryotic proteins. The *N*-linked glycosylation process occurs in eukaryotes and widely in archaea, but very rarely in bacteria. For *N*-linked oligosaccharides, a 14-sugar precursor is first added to the asparagine in the polypeptide chain of the target protein. The structure of this precursor is common to most eukaryotes, and contains 3 glucose, 9 mannose, and 2 *N*-acetylglucosamine molecules. A complex set of reactions attaches this branched chain to a carrier molecule called dolichol, and then it is transferred to the appropriate point on the polypeptide chain as it is translocated into the ER lumen.

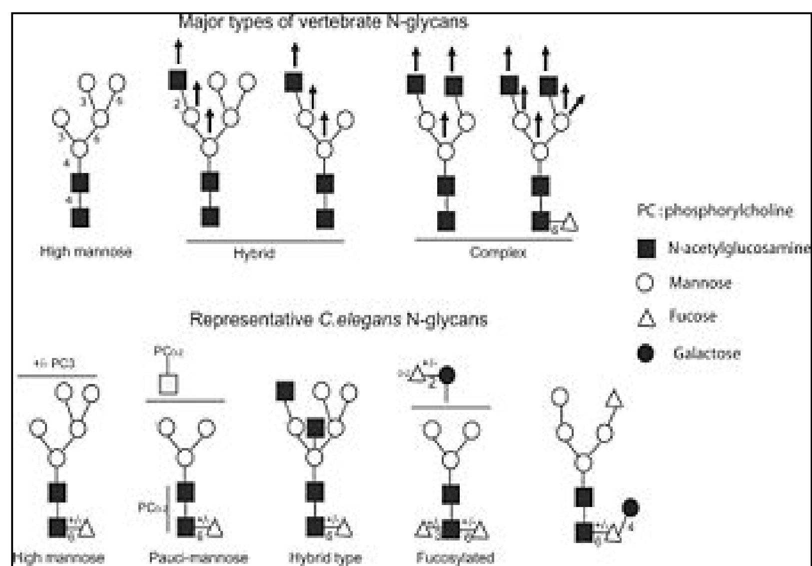


Fig. 57: Overview of the major types of vertebrate N-linked glycan glycosylation

There are two major types of *N*-linked saccharides: high-mannose oligosaccharides, in essence, just two *N*-acetylglucosamines with many mannose residues, and complex oligosaccharides, containing almost any number of the other types of saccharides, including more than the original two *N*-acetylglucosamines.

The oligosaccharide chain is attached by oligosaccharyltransferase to asparagine occurring in the tripeptide sequence Asn-X-Ser, Asn-X-Thr or Asn-X-Cys, where X could be any amino acid except Pro. This sequence is known as a glycosylation *sequon*. After attachment, once the protein is correctly folded, the three glucose residues are removed from the chain and the protein is available for export from the ER. The glycoprotein thus formed is then transported to the Golgi where removal of further mannose residues may take place. However, glycosylation itself does not seem to be as necessary for correct transport targeting of the protein, as one might think. Studies involving drugs that block certain steps in glycosylation, or mutant cells deficient in a glycosylation enzyme, still produce otherwise-structurally-normal proteins that are correctly targeted, and this interference does not seem to interfere severely with the viability of the cells. Mature glycoproteins may contain a variety of oligomannose *N*-linked oligosaccharides containing between 5 and 9

mannose residues. Further removal of mannose residues leads to a 'core' structure containing 3 mannose, and 2 *N*-acetylglucosamine residues, which may then be elongated with a variety of different monosaccharides including galactose N-acetylglucosamine, N-acetylgalactosamine, fucose and sialic acid.

B) In **O-glycosylation** (Fig. 58), the addition of sugar chains can happen on the hydroxyl oxygen on the side chain of hydroxylysine, hydroxyproline, serine or threonine. It can again be of the following types depending on the sugar moiety being attached:

a) O-N-acetylgalactosamine (O-GalNAc)- O-linked glycosylation occurs at a later stage during protein processing, probably in the Golgi apparatus. This is the addition of N-acetyl-galactosamine to serine or threonine residues by the enzyme UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase (EC 2.4.1.41), followed by other carbohydrates (such as galactose and sialic acid). This process is important for certain types of proteins such as proteoglycans, which involves the addition of glycosaminoglycan chains to an initially unglycosylated "proteoglycan core protein." These additions are usually serine O-linked glycoproteins, which seem to have one of two main functions. One function involves secretion to form components of the extracellular matrix, adhering one cell to another by interactions between the large sugar complexes of proteoglycans. The other main function is to act as a component of mucosal secretions, and it is the high concentration of carbohydrates that tends to give mucus its "slimy" feel. Proteins that circulate in the blood are not normally O-glycosylated, with the exception of IgA1 and IgD (two types of antibody) and C1-inhibitor.

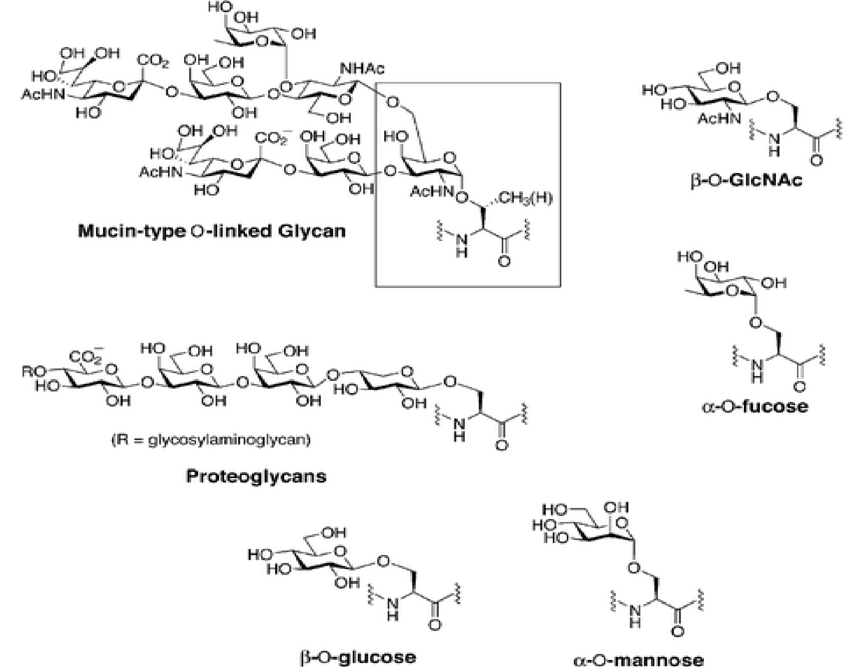


Fig. 58: Overview of the major O-linked glycan types glycosylation

b) O-fucose - O-fucose is added between the second and third conserved cysteines of EGF-like repeats in the Notch protein, and possibly other substrates by GDP-fucose protein O-fucosyltransferase 1, and to Thrombospondin repeats by GDP-fucose protein O-

fucosyltransferase 2. In the case of EGF-like repeats, the O-fucose may be further elongated to a tetrasaccharide by sequential addition of N-acetylglucoasamine (GlcNAc), galactose and sialic acids, and for Thrombospondin repeats, may be elongated to a disaccharide by the addition of glucose. Both of these fucosyltransferases have been localized to the endoplasmic reticulum, which is unusual for glycosyltransferases, most of which function in the Golgi apparatus.

c) O-glucose - O-glucose is added between the first and second conserved cysteines of EGF-like repeats in the Notch protein, and possibly other substrates by an unidentified O-glucosyltransferase.

d) O-N-acetylglucosamine (O-GlcNAc) - O-GlcNAc is added to serines or threonines by O-GlcNAc transferase. O-GlcNAc appears to occur on serines and threonines that would otherwise be phosphorylated by serine/threonine kinases. Thus, if phosphorylation occurs, O-GlcNAc does not, and vice versa. This is an incredibly important finding because phosphorylation/dephosphorylation has become a scientific paradigm for the regulation of signaling within cells. A massive amount of cancer research is focused on phosphorylation. Ignoring the involvement of this form of glycosylation, which clearly appears to act in concert with phosphorylation, means that a lot of current research is missing at least half of the picture. O-GlcNAc addition and removal also appear to be key regulators of the pathways that are deregulated in diabetes mellitus. The gene encoding the O-GlcNAc removal enzyme has been linked to non-insulin dependent diabetes mellitus. It is the terminal step in a nutrient-sensing hexosamine signaling pathway.

e) GPI anchor - A special form of glycosylation is the GPI anchor. This form of glycosylation functions to attach a protein to a hydrophobic lipid anchor, via a glycan chain.

f) C-mannosylation – A mannose sugar is added to tryptophan residues in Thrombospondin repeats. This is an unusual modification both because the sugar is linked to a carbon rather than a reactive atom like a nitrogen or oxygen and because the sugar is linked to a tryptophan residue rathen than an asparagines or serine/threonine.

Functions of Glycoproteins

In vivo, glycan moieties of glycoproteins has been known to play important roles in - maintenance of protein conformation and solubility, - proteolytic processing and stabilization of the polypeptide, - mediation of biological activity, - intracellular sorting and externalization of glycoproteins, and – embryonic development and differentiation. Glycosylation on protein residues confers on them a host of physiological functions, which are vital for maintenance of life processes. Some of these are enlisted below:

***Structural*:** Glycoproteins are found throughout matrices. They act as receptors on cell surfaces that bring other cells and proteins (collagen) together giving strength and support to a matrix (Ivatt). Proteoglycan-linking glycoproteins cross links proteoglycan molecules and is involved in the formation of the ordered structure within cartilage tissue. In nerve tissue glycoproteins are abundant in gray matter and appear to be associated with synaptosomes, axons, and microsomes. Prothrombin, thrombin, and fibrinogen are all glycoproteins that play an intricate role in the blood clotting mechanism (Gottschalk). In certain bacteria the slime layer that surrounds the outermost components of cell walls are made up of glycoproteins of high molecular weight. In addition to forming these s-layers,

glycoproteins also function as bacterial flagella. These are made up of bundles of glycoproteins protruding from the cell's surface. Their rotation provides propulsion. In plants, glycoproteins have roles in cell wall formation, tissue differentiation, embryogenesis, and sexual adhesion (certain algal species).

Protection**:** High molecular weight polymers called mucins are found on internal epithelial surfaces. They form a highly viscous gel that protects epithelium form chemical, physical, and microbial disturbances. Examples of mucin sites are the human digestive tract, urinary tract, and respiratory tracts. "Cervical mucin" is a glycoprotein found in the cervix of animals that regulates access of spermatozoa to the upper reproductive tract. Mucins are also found on the outer body surfaces of fish to protect the skin. Mucins also acts as a lubricant. Human lacrimal glands produce a glycoprotein, which protects the corneal epithelium from desiccation and foreign particles. Human sweat glands secrete glycoproteins, which protect the skin from the other excretory products that could harm the skin.

Reproduction: Glycoproteins found on the surface of spermatozoa appear to increase a sperm cell's attraction for the egg by altering the electrophoretic mobility of the plasma membrane. Actual binding of the sperm cell to the egg is mediated by glycoproteins serving as receptors on the surface of each the two membranes. The zona pellucida is an envelope made of glycoprotein that surrounds the egg and prevents polyspermy from occurring after the first sperm cell has penetrated the egg's plasma membrane. Hen ovalbumin is a glycoprotein found in egg white that serves as a food storage unit for the embryo.

Adhesion**:** Glycoproteins serve to adhere cells to cells and cells to substratum. Cell-cell adhesion is the basis for the development of functional tissues in the body. The interactions between cells are mediated by the glycoproteins on those cell's surfaces. In different domains of the body, different glycoproteins act to unite cells. For example, nerve cells recognize and bind to one another via the glycoprotein N-CAM (nerve cell adhesion molecule). N-CAM is also found on muscle cells indicating a role in the formation of myoneural junctions. With cell-substratum adhesion, glycoproteins serve as cell surface receptors for certain adhesion ligands that mediate and coordinate the interaction of cells. Substrates with the appropriate receptor will bind to the cell related to that receptor. For example, a substrate containing the glycoprotein fibronectin will be recognized and adhered to by fibroblasts. The fibroblasts will then secrete adhesion molecules and continue to spread, producing a pericellular matrix.

Hormones**:** There are many glycoproteins that function as hormones such as human chorionic gonadotropin (HCG), which is present in human pregnancy urine. Another example is erythropoietin, which regulates erythrocyte production.

Enzymes**:** Glycoprotein enzymes are of three types. These are oxidoreductases, transferases, and hydrolases. Majority of fungal enzymes are glycosylated.

Carriers**:** Glycoproteins can bind to certain molecules and serve as vehicles of transport. They can bind to vitamins, hormones, cations, and other substances.

Inhibitors**:** Many glycoproteins in blood plasma have shown antiproteolytic activity. For example, the glycoprotein a1-antichymotrypsin inhibits chymotrypsin.

Defense: In beetles pygidial glands secrete a glycoprotein disinfecting paste that covers the body and hardens. This shell provides protection against attack by bacteria and fungi (Gottschalk).

Vision: In bovine visual pigment a glycoprotein forms the outer membranes of retinal rods.

Immunological: The interaction of blood group substances with antibodies is determined by the glycoproteins on erythrocytes. Adding or removing just one monosaccharide from a blood group structure, the antigenicity and therefore a person's blood type can be altered. Many immunoglobulins are actually glycoproteins. Soluble immune mediators such as helper, suppressor, and activator cell have been shown to bind to glycoproteins found on the surface of their target cells. B and T cells contain surface glycoproteins that attract bacteria to these sites and bind them. In much the same manner, glycoproteins can direct phagocytosis. Because the HIV virus recognizes the receptor protein CD4, it binds to helper T cells that contain it.

Table: Functions of glycoproteins

Function	Glycoproteins
Structural molecule	Collagens
Lubricant and protective agent	Mucins
Transport molecule	Transferrin, ceruloplasmin
Immunologic molecule	Immunoglobins histocompatibility antigens
Hormone	Chorionic gonadotropin thyroid-stimulating hormone (TSH)
Enzyme	Various, e.g. alkaline phosphatase, cellulases, xylanases, amylases
Cell attachment-recognition site	Various proteins involved in cell-cell (sperm-oocyte), virus-cell, bacterium-cell, and hormone cell interactions
Interact with specific carbohydrates	Lectins, selectins (cell adhesion lectins), antibodies
Receptor	Various proteins involved in hormone and drug action
Affect folding of certain proteins	Calnexin, Calreticuluin
Regulation of development	Notch and its analogs, key proteins in development
Homostasis and thrmbosis	Specific glycoproteins on the surface membranes of platelets

Other specialized functions

As an antifreeze: The blood serums of Antarctic fishes freeze at -2°C, which is approximately 1°C below the melting points of their serums. This thermal hysteresis is due to the influence of some serum glycoproteins designated as antifreeze glycoproteins. They were first discovered by Arthur L. DeVries in the 1960's The temperatures of freezing and melting of aqueous solutions of the purified glycoproteins suggest that this thermal hysteresis results from the adsorption of the glycoprotein molecule onto the surface of ice

crystals. Antifreeze proteins are found in some fish, insects and plants. They bind to ice crystals and prevent them from growing to a size where they would damage the host. Specific hydrogen bonds form on the surface where protein meets ice and inhibits crystal growth. The antifreeze molecules accumulate at the interface between ice and water, not at the interface between ice and a vacuum. So it is hypothesised that a hydrophobic reaction between the protein and the neighboring water prevents the water from forming ice crystals.

***As an anticancer agent*:** Normal Gc protein (also called vitamin D binding protein), an abundant glyco-protein found in human blood serum, becomes the molecular switch to activate macrophages when it is converted to its active form, called Gc macrophage activating factor (Gc-MAF). Gc protein is normally activated by conversion to Gc-MAF with the help of the B and T cells (bone marrow-made and thymus gland-made white blood cells). Cancer cells secrete an enzyme known as alpha-N-acetylgalactosaminidase (also called Nagalase) that completely blocks conversion of Gc protein to Gc-MAF, preventing tumor-cell killing by the macrophages. This is the way cancer cells escape detection and destruction, by disengaging the human immune system. This also leaves cancer patients prone to infections and many then succumb to pneumonia or other infections. The once-weekly injection of minute amounts of Gc-MAF, just 100 nanograms (billionths of a gram), activate macrophages and allow the immune system to pursue cancer cells with vigor, sufficient to produce total long-term cures in humans.

***In protein folding*:** Glycoproteins also play important roles in mediating proper folding of proteins in the ER, which accounts for the observations that glycan addition to proteins in the ER is a cotranslational event. When inhibitors of ER glycosylation are added to cells, protein misfolding and aggregation are observed. The extent of misfolding depends on the particular protein and particular glycosylation sites with the protein. The polar CHO residues help promote solubility of folding intermediates, similar to the effects of many chaperone proteins. Some of the typical sugars found frequently in eukaryotic glycoproteins which assist in protein folding as also imparts solubility to the proteins include the following.

Sugar	**Type**	**Abbreviation**
Galactose	Hexose	Gal
Glucose	Hexose	Glc
Mannose	Hexose	Man
N-Acetylneuraminic acid	Sialic acid (nine C atoms)	NeuAc
Fucose	Deoxyhexose	Fuc
N-Acetylgalactosamine	Aminohexase	GalNAc
N-Acetylglucosamine	Aminohexase	GlcNac
Xylose	Pentose	Xyl

The glycan moieties of the folding glycoprotein also lead to binding of the protein to lectins in the ER, which serve as molecular chaperones. The most studied of these chaperones are involved in the calnexin-calreticulin cycle, and facilitate correct disulfide bond formation in the protein. After two glucose residues are removed by glucosidase I and II, the monoglucosylated protein binds to calnexin (CNX) and/or calreticulin (CRT), two

homologous ER lectins specific for monoglucosylated proteins. Once bound, another protein, ERp57, a molecular chaperone with a disulfide bond interacts with the protein. This protein has protein disulfide isomerase activity. If a glycoprotein has not folded completely, it is recognized by a glycoprotein glucosyltransferase, which adds glucose to it. This then promotes reentry into the calnexin/calreticulin cycle. Ideally, unfolded or misfolded proteins would be targeted from degradation and elimination from cells. The ER has evolved a system to accomplish this. Since folding occurs in the ER, to prevent misfolding and aggregation, the ER also contains chaperones and folding catalysts. Stress (such as through heat shock) stimulates ER chaperone activity. As a final defense mechanism, unfolded or aberrantly-folded proteins are degraded by the cytoplasmic proteasome complex. Nonnative forms of some proteins that "escape" this surveillance system can accumulate and result in disease (for example neurodegenerative diseases like Alzheimers and Parkinson's disease).

Synthesis & Regulation

Regulation and control of glycoproteins is not as straightforward as some might think. To understand regulation in glycoproteins, the enzymes that are involved in the biosynthesis pathway of these molecules are monitored. Control of glycoproteins can be seen through biosynthesis and degradation.

Synthesis: The protein part of the glycoprotein is formed at the ribosomes, where all proteins are synthesized on template represented by RNA and DNA. As a result, its structure can change only through the mutation of the genetic material of the cell. On the other hand, the carbohydrate component of a glycoprotein is not a product of the ribosome; it is synthesized somewhere in the cytoplasm, the exact site of synthesis has not been established. Since it is not directly genetically controlled, the oligosaccharide part shows a much greater variation.

In contrast with O-linked glycoproteins, where oligosaccharide assemble occurs on the polypeptide chain, N-linked glycoproteins assemble their oligosaccharide portions on a lipid linked intermediate, dolichol phosphate. The first step in oligosaccharide synthesis is the formation of that intermediate. In subsequent steps, sugars, the first being always N-acetylglucosamine (GlcNAc), are chain-like connected to dolichol phosphate. One more GlcNAc and three more mannose sugars are linked to the first GlcNAc to form the typical core of N-linked glycoproteins. Addition of any other sugar can be in any possible combination, according to the desired resulting functions. The specificity of the enzymes is very important in the synthesis process. Every sugar added, is catalyzed by a different enzyme. These group of enzymes are called glycosyltransferases. The first addition of GlcNAc to dolichol phosphate is catalyzed by the specific enzyme that cleave peptide bonds, and glycosidases (e.g example), enzymes that remove sugars one at a time from the end of an oligosaccharide chain. Both of these groups are contained in lysozomes. The lysosome attaches to a phagocyte, which has engulfed a substance that needs to be broken down, and releases its enzymes in it. Next, these enzymes begin their catalytic action. In degradation, ad in synthesis, the enzymes involved are very specific. After the glycoprotein is broken down, its amino acid and sugar components are either metabolized or can be used in the formation of another glycoprotein. Enzymatic degradation can provide much information about the structure of the oligosaccharide chains, as well as about the carbohydrate peptide linkage. For example, if a glycoprotein is treated with mannosidase (removes mannose), and mannose is released, one can conclude that mannose residues were located at the periphery of the molecule since glycosidases remove sugars from the end of

the oligosaccharide chain. Regulation and control is the organism's ultimate tool to monitor and adjust the production or degradation of different molecules.

Glycobiology and Glycomics

Our understanding of the synthesis and structure of glycan portions of glycoproteins has lagged behind our understanding of protein and nucleic acid structure and synthesis. Several reasons account for this:

- Carbohydrates are much more complex with more functional groups per carbon and with a much larger number of stereocenters, making chemical synthesis and structure determination more difficult.
- Carbohydrate chain synthesis is not directed by a template as is the synthesis of DNA, RNA, and proteins
- Synthesis is spread over two different organelles, and which allows great heterogeneity in main chain and branch chain synthesis, which provides heterogeneous samples for analysis

New techniques in analysis and synthesis of carbohydrate analogs and inhibitors of enzymes involved in CHO synthesis and degradation, as well as in genetic manipulations of gene for these enzymes, are revolutionizing our understanding of the function of carbohydrate groups on lipids and proteins. New methods to synthesize glycoproteins using recombinant DNA technology have been developed that allow synthesis of therapeutic human glycoproteins in yeast. Although they share many of the same synthetic steps, yeast glycoproteins are enriched in the high-mannose type, making them targets of the human immune system. Human and yeast glycoproteins synthesis produce the same mannose core in the ER. However, differences in synthesis occur in the Golgi. Human Golgi contain a-mannosidases I and II, which remove all but 3 Man residues from the final product. In yeast, however, these mannosidases appear to be missing so more Man residues are added (as many as 100). Human proteins made in yeast, therefore, contain many Man residues, which are recognized by the human immune system. Mutants of the yeast *Pichia pastoris* that localized glycoprotein synthesis proteins for mannosidase I and II, as well as other human glycoprotein synthesis genes, to the correct intracellular location while inactivating normal yeast gene, resulted in the production of human glycoproteins with the correct CHO structure in yeast. This may prove to have widespread use in the production of therapeutic human proteins

Examples

1. Glycophorin A, a glycoprotein that spans the plasma membrane ("Lipid bilayer") of human red blood cells. Each RBC has some 500,000 copies of the molecule embedded in its plasma membrane. Fifteen carbohydrate chains are "O-linked" to serine (Ser) and threonine (Thr) residues. One carbohydrate chain is "N-linked" to the asparagine (Asn) at position 26 (Fig. 59).

There are two polymorphic versions of glycophorin A, which differ only at residues 1 and 5, occur in humans. These give rise to the MN blood groups.

Glycophorin A is the most important attachment site by which the parasite *Plasmodium falciparum* invades human red blood cells.

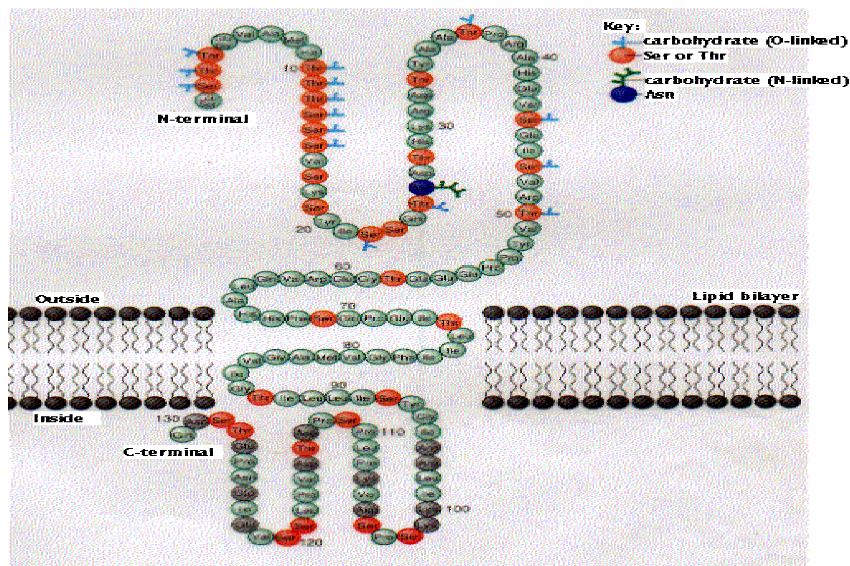


Fig. 59: Glycophorin

2. Lectins – These are sugar-binding proteins, which are highly specific for their sugar moieties. Lectins are basically non-enzymic in action and non-immune in origin. Lectins occur ubiquitously in nature. They may bind to a soluble carbohydrate or to a carbohydrate moiety that is a part of a glycoprotein or glycolipid. They typically agglutinate certain animal cells and/or precipitate glycoconjugates. They influence the biological recognition phenomena involving cells and proteins. For example, some bacteria use lectins to attach themselves to the cells of the host organism during infection.

***Biological function*:** While the function of lectins in plants is believed to be the binding of glycoproteins on the surface of parasitic cells, their role in animals also includes the binding of soluble extracellular and intercellular glycoproteins. For example, lectins found on the surface of mammalian liver cells specifically recognize galactose residues. It is believed that these cell-surface receptors are responsible for the removal of certain glycoproteins from the circulatory system. Another example is the mannose-6-phosphate receptor that recognizes hydrolytic enzymes containing this residue and subsequently targets these proteins for delivery to the lysosomes. (one defect in this particular system is known as I-cell disease.) Lectins serve many different biological functions from the regulation of cell adhesion to glycoprotein synthesis and the control of protein levels in the blood. Lectins are also known to play important roles in the immune system by recognizing carbohydrates that are found exclusively on pathogens, or that are inaccessible on host cells. Examples are the lectin complement activation pathway and Mannose binding lectin.

The function of lectins in plants is still uncertain. Once thought to be necessary for rhizobia binding, this proposed function was ruled out through lectin-knockout transgene studies. The large concentration of lectins in plant seeds decreases with growth, and suggests a role in plant germination and perhaps in the seed's survival itself.

3. Mucins – Mucins are secreted in the mucus of the respiratory and digestive tracts. The sugars attached to mucins give them considerable water-holding capacity and also make them resistant to proteolysis by digestive enzymes. Glycoproteins are important for immune cell recognition, especially in mammals. Examples of glycoproteins in the immune system are:

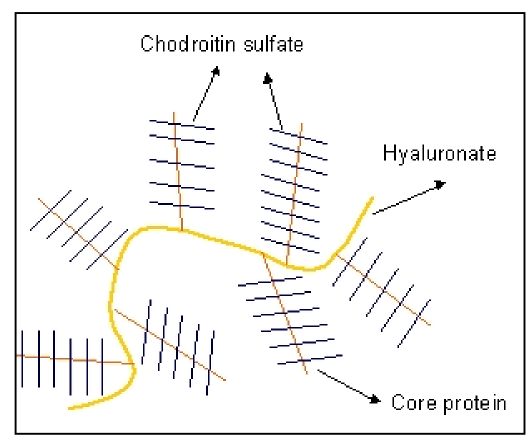
- Molecules such as antibodies (immunoglobulins), which interact directly with antigens
- Molecules of the major histocompatibility complex (or MHC), which are expressed on the surface of cells and interact with T-cells as part of the adaptive immune response.

Other examples of glycoproteins include:
- Components of the zona pellucida, which surrounds the oocyte, and is important for sperm-egg interaction.
- Structural glycoproteins, which occur in connective tissue. These help bind together the fibers, cells, and ground substance of connective tissue. They may also help components of the tissue bind to inorganic substances, such as calcium in bone.
- Enzymes, immunoglobulins etc. are glycoproteins with various applications.

Proteoglycans

Proteoglycans represent a special class of glycoproteins that are heavily glycosylated. They consist of a core protein with one or more covalently attached glycosaminoglycan chain(s). These glycosaminoglycan (GAG) chains are long, linear carbohydrate polymers that are negatively charged under physiological conditions, due to the occurrence of sulphate and uronic acid groups (Fig. 60). Proteoglycans can be categorised depending upon the nature of their glycosaminoglycan chains. These chains may be:
- chondroitin sulfate and dermatan sulfate
- heparin and heparan sulfate
- keratan sulfate

(These have been described in derivatives of polysaccharides part earlier.)



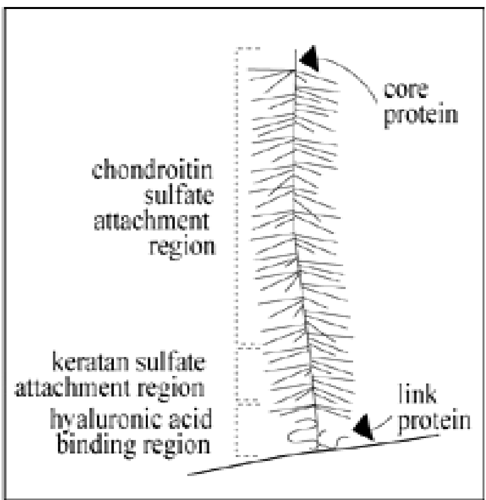


Fig. 60: Structure of proteoglycan

Proteoglycans can also be categorized by size. Examples of large proteoglycans are aggrecan, the major proteoglycan in cartilage, and versican, present in many adult tissues including blood vessels and skin. The small leucine rich repeat proteoglycans (SLRPs) include decorin, biglycan, fibromodulin and lumican.

Synthesis: The protein component of proteoglycans is synthesized by ribosomes and translocated into the lumen of the rough endoplasmic reticulum. Glycosylation of the proteoglycan occurs in the Golgi apparatus in multiple enzymatic steps. First a special link tetrasaccharide is attached to a serine side chain on the core protein to serve as a primer for polysaccharide growth. Then sugars are added one at the time by glycosyl transferase. The completed proteoglycan is then exported in secretory vesicles to the extracellular matrix of the cell.

Function: Proteoglycans are a major component of the animal extracellular matrix, the 'filler' substance existing between cells in an organism. Here they form large complexes, both to other proteoglycans, to hyaluronan and to fibrous matrix proteins (such as collagen). They are involved in binding cations (such as sodium, potassium and calcium) and water, and also regulating the movement of molecules through the matrix. Evidences shows that they can affect the activity and stability of proteins and signalling molecules within the matrix. Individual functions of proteoglycans can be attributed to either the protein core or the attached GAG chain.

Proteoglycans and disease: An inability to break down proteoglycans is characteristic of a group of genetic disorders, called mucopolysaccharidoses. The inactivity of specific lysozomal enzymes that normally degrade glycosaminoglycans leads to the accumulation of proteoglycans within cells. This leads to a variety of disease symptoms, depending upon the type of proteoglycan that is not degraded.

Structure of proteo glycans

Glycoforms are types of proteins (isoforms) with specific type of glycoprotein attached to them by either posttranslational or cotranslational modifications. In other words, two proteins would be of the same glycoform if they carried the same glycoprotein. Glycoform is defined as a subunit of molecules with identical polypeptide sequences but with different glycans present at the sites of glycosylation. It is very important to analyze the glycoform, because different glycoforms of the same glycoprotein have different biological properties. It has been estimated, for example, that recombinant tissue plasminogen activator (rtPA) may contain as many as 11,500 different glycoforms. One interesting example appears to be Glycodelin, a human secreted glycoprotein that appears in a small number of glycoforms, exhibits diverse biological activities, such as in contraception and immunosuppression. Moreover, different tissue-specific glycoforms appear to mediate diverse functions. Quite unusually, the glycodelin *N*-linked glycans differ between the male and female glycoforms. The fact that these glycans are fundamental for exerting the physiological activities of the different glycoforms makes them an interesting target for glycobiology research.

Conformation of polysaccharides

Conformation of the polysaccharides gives the idea of three-dimensional orientation of a specific polysaccharide in space and the orientation of the glycosidic groups about projections of sugar moieties in the chain.

A. Three-dimensional orientation

In order to completely describe the three dimensional orientation of a specific polysaccharide in space, we need to know the following general notations.

Direction of Numbering: The chain is numbered from the reducing glycose residue to the non-reducing glycosyl group. Thus i refers to a particular saccharide unit in the polymer chain, (i - 1) to the adjacent unit in the direction away from the non-reducing end and (i + 1) in the direction of the non-reducing end. This direction of numbering is chosen so that gain or loss of a residue at the non-reducing end by transfer of a glycosyl group does not change the numbering of every unit in the chain. Some polysaccharides lack reducing end groups and are initiated by formation of a glycosidic linkage, e.g. to O1 of another sugar, to an alditol or to an alcohol. For these polysaccharides the residue forming this glycosidic linkage is regarded as the first member of the chain.

Monomeric unit: The monomeric unit is the monosaccharide. The oxygen of the glycosidic bond is part of the residue glycosylated. Since its position is of importance in specifying the unit, the torsion angle at the glycosidic bond is included in the characteristics of the sugar residue.

Atomic numbering: The notation used conforms with that being proposed for specifying polynucleotide conformation. Atoms are thus designated C3, O2, H4, etc. The hydrogen atoms of a methylene group may be distinguished by an additional number, e.g. H61 and H62. When it is necessary to indicate the particular saccharide unit its number may be added in parenthesis, e.g. O3(i), C4(i+1), H61(i-1) (Fig. 61).

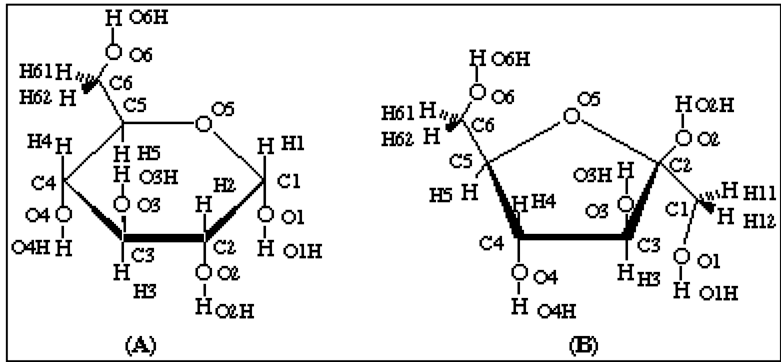


Fig. 61: Atomic numbering

Notation for atomic numbering of (A) a hexopyranose and (B) a hexofuranose unit

Interatomic distance: In tabulating interatomic distances, non-bonded atoms are represented with a single dot between them, e.g. O2.C3, and covalent bonds are represented by hyphens between atoms, e.g. C1-C2. Hydrogen bonds are denoted by dotted lines whether or not the hydrogen is shown, e.g. O6...O5 or O6-H...O5. The atom donating the hydrogen is written first if specification is possible.

Bond Angles: The bond angle included between three atoms A-B-C is written 'tau' as τ (A, B, C). If there is no ambiguity because the central atom is bivalent, this may be abbreviated to τ (B). The angle at the ring-oxygen atom of an aldopyranose may thus be written as τ (C1, O5, C5) or τ (O5).

Ring Shape: The ring shape of a sugar residue can be defined either by the endocyclic torsion angles or in terms of the notation for conformations of five and six-membered monosaccharide rings.

Endocyclic Torsion Angles: To provide a complete description of the sugar ring conformation, it is necessary to specify the endocyclic torsion angles about at least some of the ring bonds, in addition to the bond lengths and bond angles. These ring torsion angles, denoted by the symbol *υ* (nu), can be described by adding a number indicating the bond as listed in. The torsion angle of the atoms A-B-C-D is the angle between A-B and C-D in a projection of the four atoms on to a plane normal to B-C. It is considered positive when the bond to the front, viewed along the central bond, must be rotated clockwise to eclipse the bond to the rear.

Conformation of Side Groups: A ring substituent of a pyranosidic sugar unit. e.g. a hydroxyl group, may be designated as being axial or equatorial in a given conformation. For the precise specification of the orientation of a polyatomic ring substituent, it is necessary to specify the torsion angle about the exocyclic bond. The reference atom in the ring is the carbon atom with the number one lower than that of the substituted carbon, unless substitution is on the anomeric carbon, when the ring oxygen is the reference atom. The reference atom in an exocyclic-CH2X group is X. The exocyclic torsion angle is denoted by χ (chi), followed by the atoms to which it refers. *e.g.* (C1-C2-O2-H) or, if no ambiguity arises, simply by χ^2.

B. Orientation of the glycosidic groups

The structures are more precisely defined with repect to the following notations corresponding to the orientation of the glycosidic groups.

Designation of Bonds: The glycosidic bonds are part of the backbone of the polysaccharide. The glycosidic linkage is most easily described using the symbols for the monomeric units and the locants together with the anomeric descriptor, e.g. in cellobiose: Glc (i)(β 1→4) Glc (i^{-1}). If a more detailed description is necessary, this can be given in the form C1(i)-O4(i^{-1})-C4(i^{-1}). (Fig. 62)

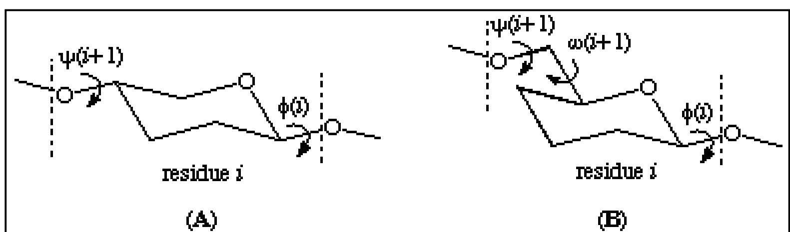


Fig. 62: Designation of bonds

Notation for the torsion angles specifying the orientation of glycosidic bonds for a pyranose unit (A) varrying a ring hydroxyl that is glycosylated and (B) in which the glycosylated hydroxyl group is on an exocyclic carbon. The limits of the i th residue are indicated by the vertical dash lines. The residue number attached to φ, ψ and ω is conventionally that of the glycosylating residue.

Torsion Angles: Two torsion angles, φ (phi) and ψ (psi), are required to describe the glycosidic bond from the i th unit to a carbon atom located in the ring of the (i^{-1}) th unit. The angle φ about the bond from the anomeric carbon to the oxygen that joins the two residues is specified using the ring oxygen as a reference atom. The torsion angle ψ about the bond from the glycosylated oxygen of the (i^{-1}) th residue to a carbon of this residue uses the carbon atom one lower in numbering as a reference atom. Since this angle relates to the mode of attachment of the ith residue, it may be designated ψ (i). When the glycosidic bond does not involve a carbon atom located in the ring, but rather on a side chain, the angle ω (omega), around the next C-C bond is also of importance. For a 1→6 linked aldohexopyranose, this angle is the exocyclic angle c5 of the (i^{-1}) th residue. Nevertheless it may be designated as ω (i) since it refers to attachment of the i th residue. The torsion angle is that of the atoms OX-CX-C $\overline{X-1}$ -C $\overline{X-2}$. where X is the number of the carbon atom whose hydroxyl group if glycosylated.

Helix Characteristics: In the description of helices or helical segments the following symbols may be used:
n = number of repeating units per turn
h = unit height (translation per repeating unit along the helix axis)
t = 360deg./n = unit twist (angle of rotation per repeating unit about the helix axis)
p = pitch height of helix = n . h.
Note. The repeating unit in a homopolysaccharide is a sugar residue. Heteropolysaccharides may possess repeating units of two or more residues.

Rare sugars

Rare sugars are defined by the International Society of Rare Sugars (ISRS) as monosaccharides and their derivatives that are rare in nature. They are hardly available for research purposes because of their expensiveness. "Izumoring", a structural framework containing all 34 six-carbon monosaccharides linked by enzymatic reactions, has been proposed following the discovery of a key enzyme that converts abundantly occurring monosaccharides in nature into rare sugars. This has made possible the mass production of rare sugars from inexpensive sugars such as D-glucose or D-fructose.

Rare Sugars are mostly used in pharmaceuticals as precursors for a wide variety of carbohydrate-based drugs. These include nucleoside analogues, which are used in antiviral applications such as HIV, HBV and HCV. Another important class of compounds is complex oligosaccharides and olignonucleotides, which may be used as anti-inflammatory or anti-cancer agents, as well as in highly specific chronic pain relievers. They are also being used as precursors in the production of flavor chemicals, such as natural furanones and Maillard reaction savoury flavours. Furthermore, some of the rare sugars products have applications as nutraceuticals or they may be used in high-end cosmetic products. They are enlisted under the D & L series depending on their chirality.

D series		L series	
D- Sorbose		L- Sorbose	
		L- Fructose	
D- Psicose		L- Psicose	
D- Tagatose		L- Tagatose	
D- Gulose	D- Gulitol	L- Gulose	L- Gulitol
D- Idose	D- Iditol	L- Idose	L- Iditol
D- Talose	D- Talitol	L- Talose	L- Talitol
	D- Galactitol	L- Galactose	L- Galactitol
	D- Mannitol	L- Mannose	L- Mannitol
	D- Glucitol	L- Glucose	L- Glucitol
D- Altrose	D- Altritol	L- Altrose	L- Altritol
D- Allose	D- Allitol	L- Allose	L- Allitol

Carbohydrate assays

Carbohydrates can be assayed using the methods below

Molisch Test (Carbohydrates): The **Molisch test** is a general **test** for the presence of carbohydrates. **Molisch** reagent is a solution of alpha-naphthol in 95% ethanol. This **test** is useful for identifying any compound, which can be dehydrated to furfural or hydroxymethylfurfural in the presence of H_2SO_4. Furfural is derived from the dehydration of pentoses and pentosans, while hydroxymethylfurfural is produced from hexoses and hexosans. Oligosaccharides and polysaccharides are hydrolyzed to yield their repeating monomers by the acid. The alpha-naphthol reacts with the cyclic aldehydes to form purple colored condensation products. Although this **test** will detect compounds other than carbohydrates (i.e. glycoproteins), a negative result indicates the **ABSENCE** of carbohydrates.
[Prepare Molisch's reagent by dissolving 0.5 g reagent grade α-naphthol in10 mL of 95% ethanol. Store the reagent, protected from light, at room temperature. To test for carbohydrates, add 0.02 mL of the reagent to 1 mL of 0.1% carbohydrate (1 mg/mL) solution in a small test tube. After mixing, tilt the tube and carefully add without mixing, 0.5 mL of concentrated sulfuric acid by pouring it down the side of the tube. A red-violet layer at the interface between the acid (bottom) and aqueous (upper) layers is a positive test for carbohydrates]

Bial's Test (Pentoses): Bial's reagent uses orcinol, HCl, and $FeCl_3$. Orcinol forms colored condensation products with furfural generated by the dehydration of pentoses and pentosans. It is necessary to keep the concentration of the sugar solution around 0.2M by appropriate dilutions in this test.
[Prepare Bial's reagent by dissolving 0.3 g reagent-grade orcinol and 0.05 g ferric chloride in 100 mL of concentrated (12 M) HCl. Store the reagent protected from light. To test for pentoses, add 0.05 mL of 0.1% carbohydrate solution in water to 1 mL of Bial's reagent, and heat the solution in a boiling water bath for 2 minutes. * A blue-green color indicates

pentoses or nucleotides containing pentoses; a yellow-green color indicates hexoses, and disaccharides are yellow.]

Resorcinol (Seliwanoff's) Test (Ketohexoses): ***Seliwanoff's reagent*** contains resorcinol in 6 M hydrochloric acid. Hexoses undergo dehydration when heated in this reagent to form hydroxymethylfurfural, which condenses with resorcinol to give a red product. Ketohexoses (such as fructose) and disaccharides containing a ketohexose (such as sucrose) form a cherry-red condensation product. Other sugars may produce yellow to faint pink colors.
[Prepare Seliwanoff's reagent by dissolving 0.05 g of reagent-grade resorcinol in 100 mL of 3 M HCl. Store protected from light. To test for ketohexoses, add 0.1 mL of a 1% carbohydrate solution in water to 1 mL of the reagent, and heat the solution in a boiling water bath for 5 minutes.* A deep red colored precipitate within 5 minute indicates ketohexoses. Sucrose gives a positive ketohexose test because of partial hydrolysis to glucose and fructose. Other sugars give a red color upon prolonged heating.]

Benedict's Test (Reducing sugars): Alkaline solutions of copper are reduced by sugars having a free aldehyde or ketone group, with the formation of colored cuprous oxide. Benedict's solution is composed of copper sulfate, sodium carbonate, and sodium citrate (pH 10.5). The citrate will form soluble complex ions with Cu^{++}, preventing the precipitation of $CuCO_3$ in alkaline solutions.
[Dissolve 1.73 g trisodium citrate (dihydrate) and 1.0 g anhydrous sodium carbonate in 8 mL of warm distil H_2O (Solution A). Dissolve copper sulfate (pentahydrate) (1.73 g) separately in 20 mL of distil H_2O (Solution B). Immediately before using, prepare Benedict's reagent by mixing 0.8 mL of Solution A with 0.2 mL of Solution B. To test for reducing sugars, add 0.2 mL of a 1% carbohydrate solution to 1 mL of Benedict's reagent and heat in a boiling water bath for 5 minutes. A brick-red precipitate indicates a positive test for reducing sugars.]

Barfoed's Test (Reducing monosaccharides): This reaction detects reducing monosaccharides in the presence of disaccharides. This reagent uses copper ions to detect reducing sugars in an acidic solution. Barfoed's reagent is copper acetate in dilute acetic acid (pH 4.6). The color changes are the same as that in Benedict's test.
[Prepare Barfoed's reagent by dissolving 0.66 g cupric acetate (monohydrate) and 0.18 mL glacial acetic acid in 10 mL of distil H_2O. To test for reducing monosaccharides, add 0.3 mL of 1% carbohydrate solution to 0.6 mL of Barfoed's reagent and heat in a boiling-water bath for 5 minutes,* then cool to room temperature. A copious amount of brick-red precipitate indicates a reducing monosaccharide. Some hydrolysis of disaccharides may lead to trace precipitates.]

Iodine Test (Starch/Amylose): The use of Lugol's iodine reagent (IKI) is useful to distinguish starch and glycogen from other polysaccharides. Lugol's iodine yields a blue-black color in the presence of starch. Glycogen reacts with Lugol's reagent to give a brown-blue color. Other polysaccharides and monosaccharides yield no color change; the **test** solution remains the characteristic brown-yellow of the reagent. It is thought that starch and glycogen form helical coils. Iodine atoms can then fit into the helices to form a starch-iodine or glycogen-iodine complex. Starch in the form of amylose and amylopectin has less branches than glycogen. This means that the helices of starch are longer than glycogen, therefore binding more iodine atoms. The result is that the color produced by a starch-iodine complex is more intense than that obtained with a glycogen-iodine complex.
[A few drops of 0.01 M iodine in 0.12 M KI are added to a 1% solution of the carbohydrate in question. The immediate formation of a vivid blue color indicates amylose.]

Glucose Test (GOD-PAP): This method involves colorimetric determination of glucose by employing two successive enzymatic reactions; Glucose is first oxidized to gluconic acid by glucose oxidase with the concomitant production of hydrogen peroxide. This is subsequently converted to a quinone by peroxidase. The red quinone generated is then measured colorimetrically to determine the amount of glucose present initially
[Prepare the assay reagent by dissolving a capsule (Sigma 510-6) containing 500 units of glucose oxidase and 100 units of horseradish peroxidase with buffer salts in 100 mL of distilled water, and adding 1.6 mL of 2.5 mg/mL o-dianisidine dihydrochloride (Sigma 510-50). Combine 0.1 mL of a sample solution containing 0.25 to 3 mg/mL of carbohydrate with 1 mL of the assay reagent. A brown color within 30 minutes indicates glucose. * Heat in a fume hood.
Precaution:
1. Use a glass Pasteur pipette to add the H_2SO_4.
2. Do not use a mechanical pipettor with concentrated acids.

Reminder: Always add acid to water.]

Mucic Acid Test for Galactose: Oxidation of most monosaccharides by nitric acid yields soluble dicarboxylic acids. However, oxidation of galactose yields an insoluble mucic acid. Lactose also yields a mucic acid, due to hydrolysis of the glycosidic linkage between its glucose and galactose subunits.

Summary of carbohydrate assay

Sugar	Molisch	Bial	$(\phi)_2$NH	G.O.[§]	Resorcinol	Benedict	Barfoed	Iodine
Ribose	+ (r/v)	(bl/gr)	-	-	-	+(r ppt)	+ (r ppt)	-
DNA*	+ (r/v)	(y)	(bl/gr)	-	-	+(r ppt)	+ (r ppt)	-
Fructose	+ (r/v)	(y/gr)	-	-	+ (r)	+(r ppt)	+ (r ppt)	-
Galactose	+ (r/v)	(y/gr)	-	-	-	+(r ppt)	+ (r ppt)	-
Glucose	+ (r/v)	(y/gr)	-	+(br)	-	+(r ppt)	+ (r ppt)	-
Sucrose	+ (r/v)	(y)	-	-	+ (r)	-	-	-
Lactose	+ (r/v)	(y)	-	-	-	+(r ppt)	-	-
Maltose	+ (r/v)	(y)	-	-	-	+(r ppt)	-	-
Amylose	+ (r/v)	(y)	-	-	-	-	-	+ (bl)
Glycogen	+ (r/v)	(y)	-	-	-	-	-	+ (br)

+ = positive; - = negative; bl = blue; br = brown; gr = green; r = red; v = violet; y = yellow; ppt = precipitate
* To test for DNA, the sample is first hydrolyzed in 10% trichloroacetic acid at 95°C for 10 minutes then diluted with two volumes of water before assaying with the diphenylamine test.
[§] **GO = glucose oxidase**

Suggested Reading

1. Morrison, R.T. & Boyd, R.N. (1992) *Organic Chemistry*, 6th edn, Benjamin Cummings, San Francisco. Chapters 34 and 35 cover the structure, stereochemistry, nomenclature, and chemical reactions of carbohydrates.
2. Pigman, W. & Horton, D. (Eds) (1970, 1972, 1980) the *Carbohydrates: Chemistry and Biochemistry, Vols IA, IB, IIA, and IIB,* Academic Press, Inc., New York. Comprehensive treatise on carbohydrate chemistry.
3. Fukuda, M. & Hindsgaul, O. (1994) *Molecular Glycobiology*, IRL Press at Oxford University Press, Inc., New York. Thorough, advanced treatment of the chemistry and biology of cell surface

carbohydrates. Good chapters on lectins, carbohydrate recognition in cell-cell interactions, and chemical synthesis of oligosaccharides.

4. David L. Nelson. Michael M. Cox. *Principles of Biochemistry*; 4th edition, 2006, W. H. Freeman and Company. NY. 2006.
5. KleinKauf . Dohren. Jaenicke. *The roots of modern biochemistry*. Waller De Gruyter Berlin. NY. 1988.
6. Stryer Lubert. *Biochemistry.* 4th edition. 1995. W. H. Freeman and Company. NY.
7. Donald Voet & Judith G Voet. *Principles of Biochemistry.* 4th edition. 2004. J Wiley & Sons. NY.
8. Brooks SA, Dwek, MV, Schumacher. *Functional and Molecular Glycobiology*. (2002). U. Bios Scientific Publishers.
9. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD, *Molecular Biology of the Cell,* (3rd Edition). (1996), W. H. Freeman and Company
10. Benjamin Cummings, *Biology* (4th edition) Campbell, N. A. (1996), New Work.
11. *Merck Index,* (12th Edition), (1996) Merck Research Laboratories Division of Merck & Co. Inc.
12. Sharon, N., Lis, H. *Lectins,* 2nd Edition (2003) publishers Kluwer Academic.
13. Encyclopedia of Biological Chemistry (Lennarz & Lane, EDs.) (2004) Academic Press/Elsevier, Oxford,
14. Iozzo, M, V. *Proteoglycans: structure, biology and molecular interactions.* (2000). Marcel Dekker Inc. New York.
15. www.halosource.com
16. www.wikiepedia.com
17. Qiyong P Liu, Gerlind Sulzenbacher, Huaiping Yuan1, Eric P Bennett, Greg Pietz, Kristen Saunders, Jean Spence, Edward Nudelman, Steven B Levery, Thayer White, John M Neveu, William S Lane, Yves Bourne, Martin L Olsson, Bernard Henrissat & Henrik Clausen. Bacterial glycosidases for the production of universal red blood cells; 2007 Nature biotechnology; Edward Nudelman. Nature Biotechnology **25**, 454 - 464 (2007)
18. Pei Zhang, Scott Snyder, Peter Feng, Parastoo Azadi, Shusheng Zhang, Silvia Bulgheresi, Kenneth E. Sanderson, Johnny He, John Klena, and Tie Chen; Role of N-Acetylglucosamine within Core Lipopolysaccharide of Several Species of Gram-Negative Bacteria in Targeting the DC-SIGN (CD209). The Journal of Immunology, **177**: 4002-4011(2006).
19. Francisco J. Alvarez* and James B. Konopka. Identification of an N-Acetylglucosamine Transporter That Mediates Hyphal Induction in Candida albicans. Mol Biol Cell. March; **18**(3): 965–975. (2007)
20. Wells L, Vosseller K, Hart GW; A role for N-acetylglucosamine as a nutrient sensor and mediator of insulin resistance. Cellular and Molecular Life Sciences (CMLS) **Volume 60**, Number 2 / February, (2003)
21. Roseman S. "Reflections on glycobiology". J. Biol. Chem. **276** (45): 41527-42(2001).
22. Anderson JW, Nicolosi RJ, Borzelleca JF. Glucosamine effects in humans: a review of effects on glucose metabolism, side effects, safety considerations and efficacy. Food and Chemical Toxicology.; **43** (2): 187- 201.(2005)
23. Mireia Garriga-Canut, Barry Schoenike, Romena Qazi, Karen Bergendahl, Timothy J Daley, Rebecca M Pfender, John F Morrison, Jeffrey Ockuly, Carl Stafstrom, Thomas Sutula & Avtar Roopra, "2-Deoxy-D-glucose reduces epilepsy progression by NRSF-CtBP–dependent metabolic regulation of chromatin structure", Nature Neuroscience, **9**, 1382 - 1387 (2006).
24. Funderburgh JL. "Keratan sulfate: structure, biosynthesis, and function". Glycobiology **10** (10): 951-958, (2000).
25. Distribution of proteoglycans antigenically related to corneal keratan sulfate proteoglycan". J. Biol. Chem. **262** (24), (1987).
26. Iozzo, R. V. Matrix proteoglycans: from molecular design to cellular function. Annu. Rev. Biochem. **(67)**: 609-652. (1998).
27. Stern R "Hyaluronan catabolism: a new metabolic pathway". Eur J Cell Biol **83** (7): 317-25. (August 2004).
28. W. Boerjan, J. Ralph, M. Baucher "Lignin bios". Ann. Rev. Plant Biol. **54**: 519-549, (Jun 2003).
29. Davin, L.B.; Lewis, N.G. "Lignin primary structures and dirigent sites". Current Opinion in Biotechnology **16**: 407-415, (2005).

View publication stats