CARBOHYDRATES AND GLYCOBIOLOGY

Dr.Sulieman Al-Khalil

Carbohydrates: Earth's Most Abundant Biomolecules

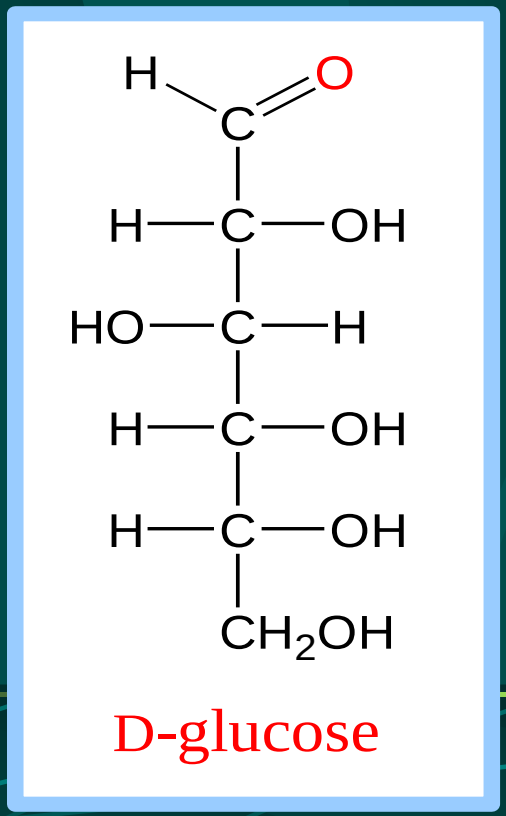
Carbohydrates (glycans) have the following basic composition:

$(CH_2O)_n$ or $H-\overset{\overset{|}{}}{\underset{\underset{|}{}}{C}}-OH$

- Monosaccharides - simple sugars with multiple OH groups. Based on number of carbons (3, 4, 5, 6), a monosaccharide is a triose, tetrose, pentose or hexose.
- Disaccharides - 2 monosaccharides covalently linked.
- Oligosaccharides - a few monosaccharides covalently linked.
- Polysaccharides - polymers consisting of chains of monosaccharide or disaccharide units.

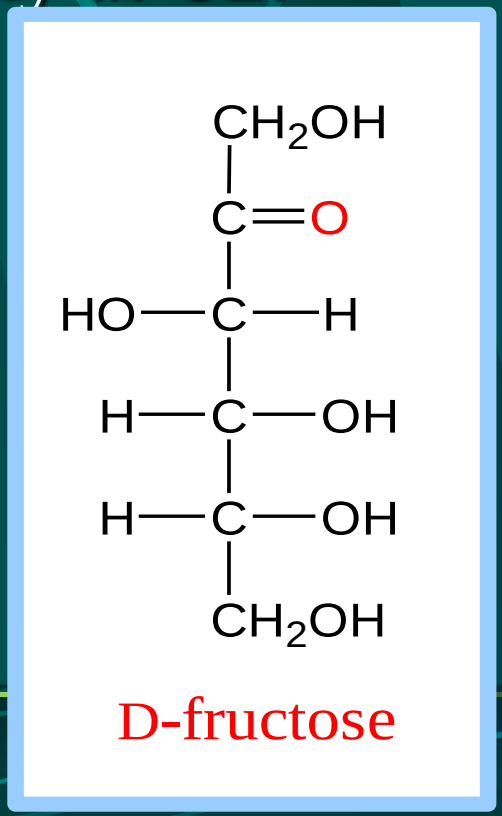
Monosaccharides

Aldoses (e.g., glucose) have an aldehyde group at one end.



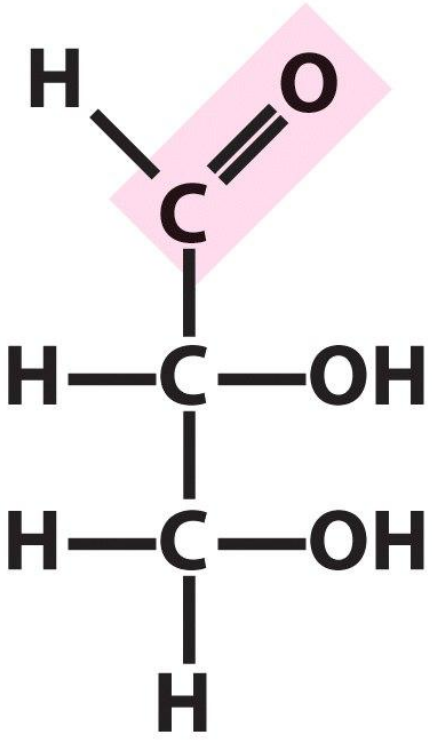
D-glucose

Ketoses (e.g., fructose) have a keto group, usually at C2.

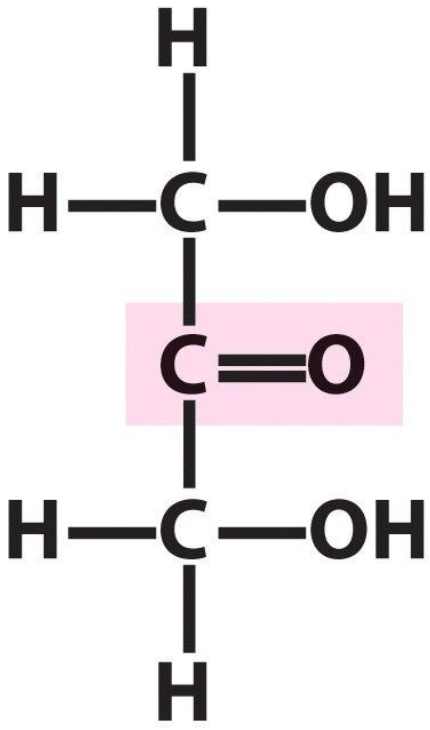


D-fructose

Trioses: The Simplest Monosaccharides

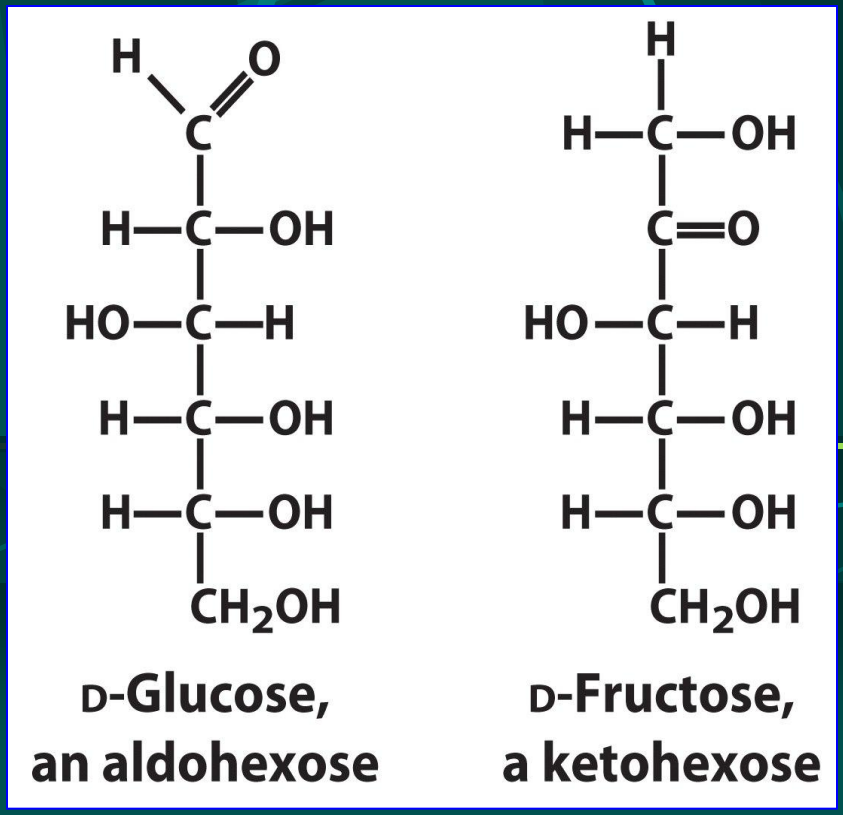


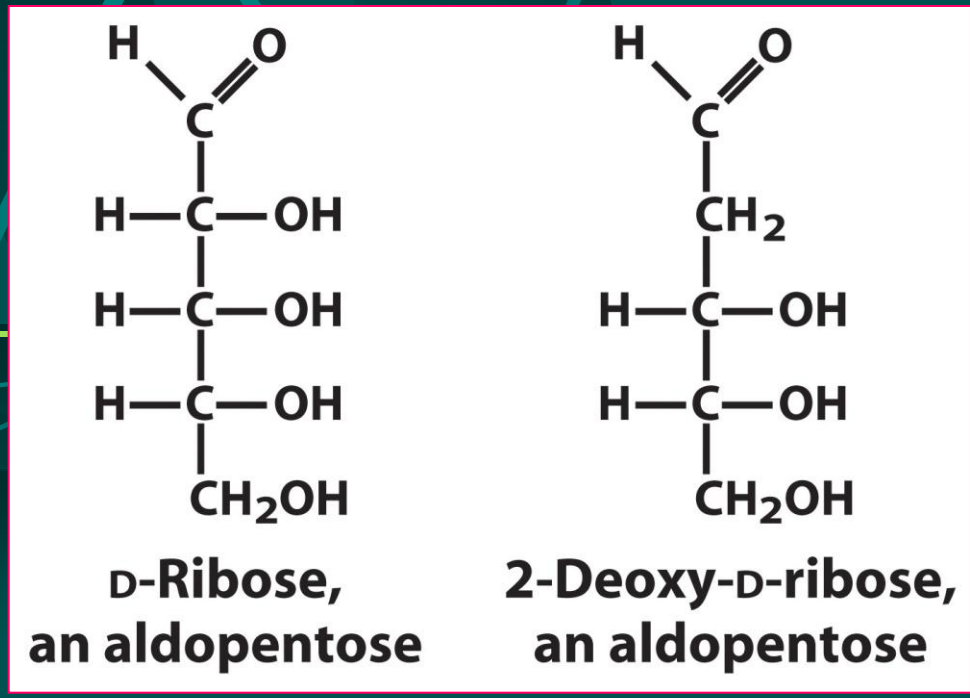
Glyceraldehyde, an aldotriose



Dihydroxyacetone, a ketotriose

Hexoses and Pentoses

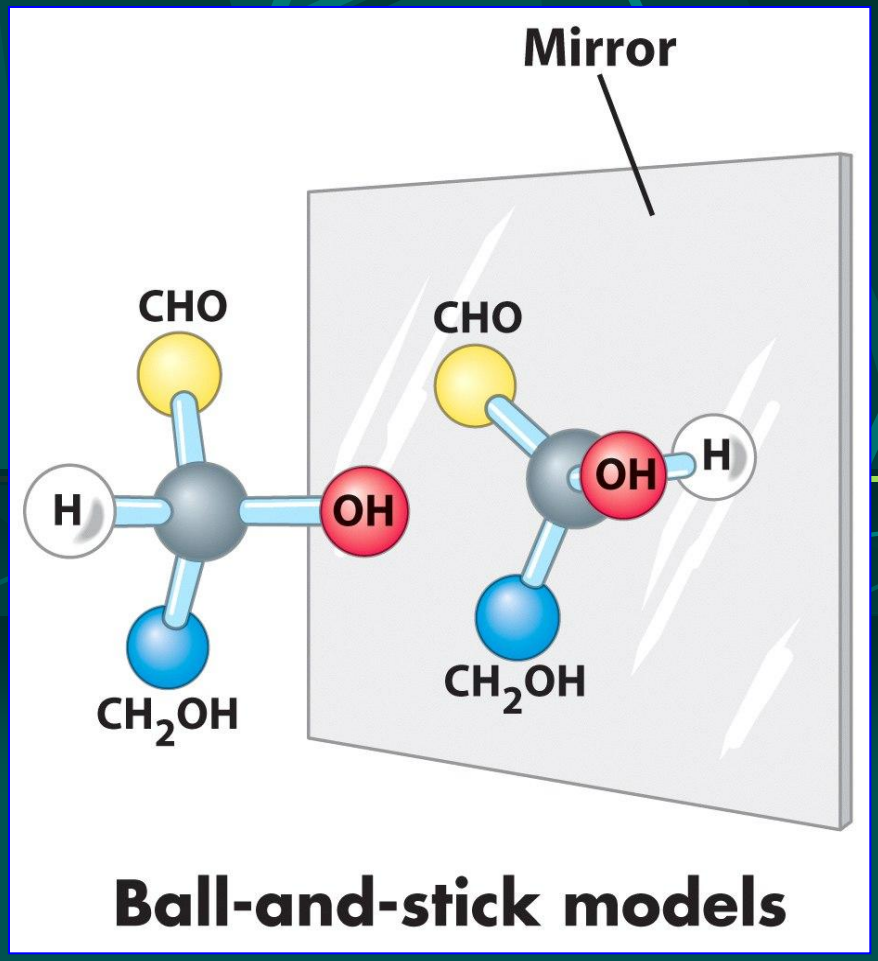


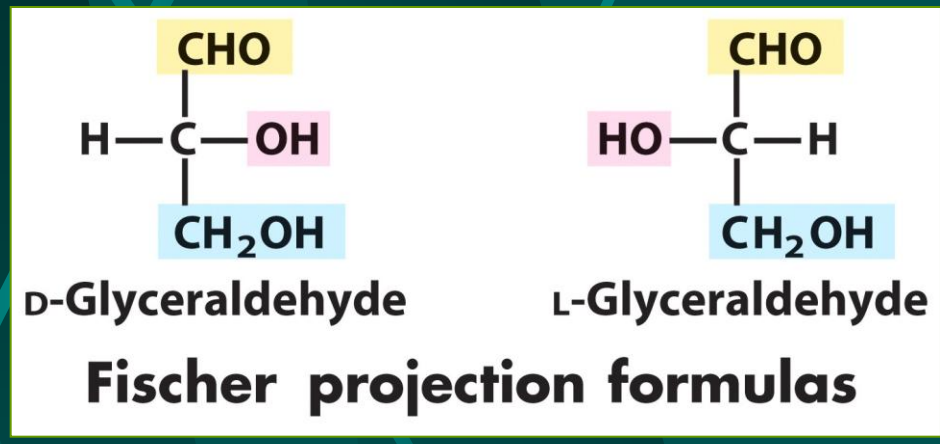


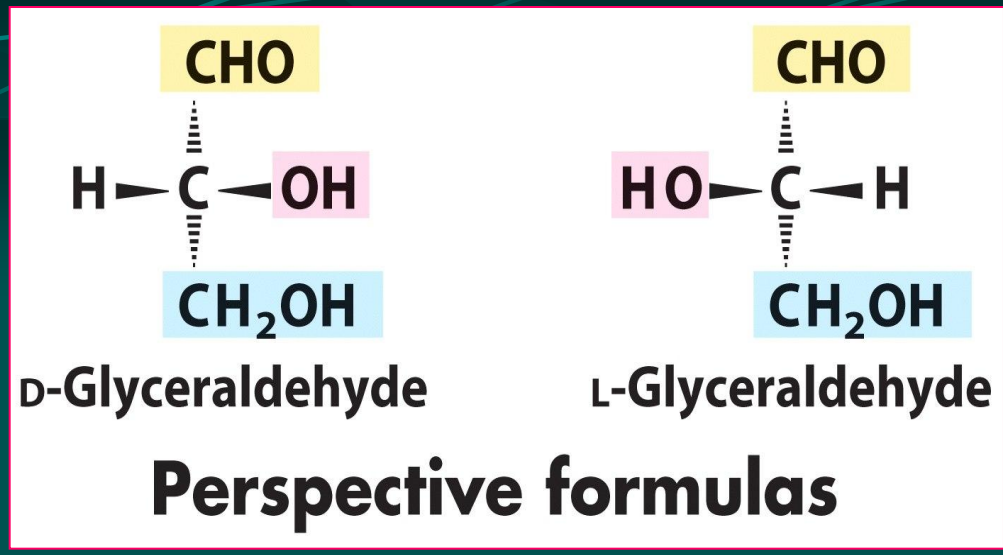
Hexoses: most abundant

Pentoses: RNA and DNA

Stereoisomerism of Monosaccharides



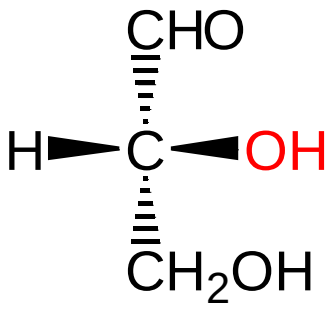




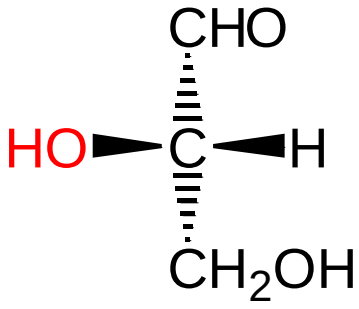
D vs L Designation

D & L designations are based on the configuration about the single asymmetric C in glyceraldehyde.

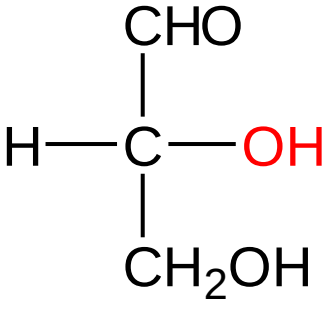
The lower representations are Fischer Projections.



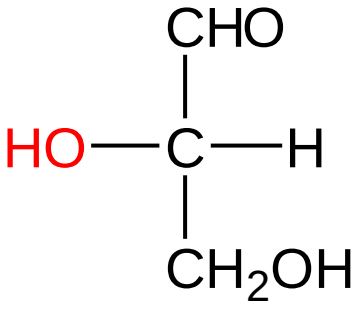
D-glyceraldehyde



L-glyceraldehyde



D-glyceraldehyde

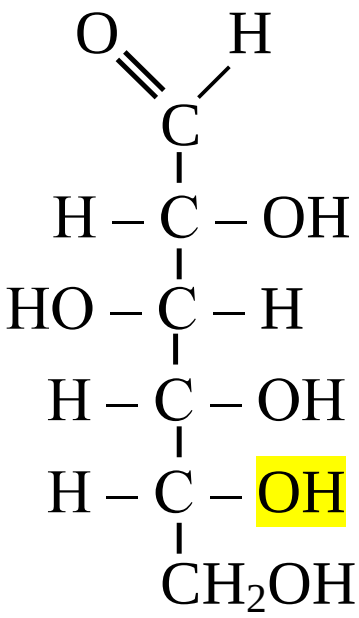


L-glyceraldehyde

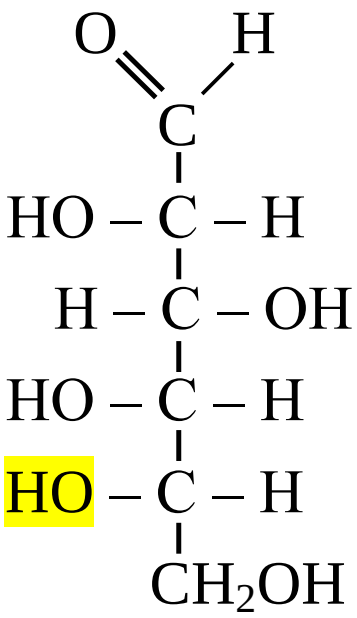
Sugar Nomenclature

For sugars with more than one chiral center, **D** or **L** refers to the asymmetric **C** farthest from the aldehyde or keto group.

Most naturally occurring sugars are D isomers.



D-glucose

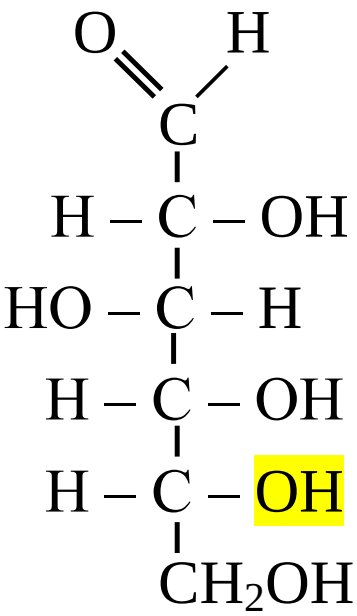


L-glucose

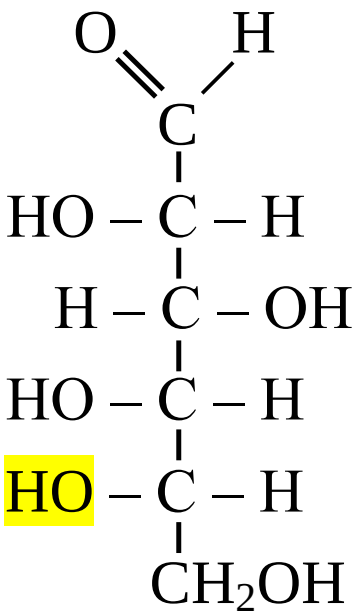
D & L sugars are mirror images of one another.

They have the same name, e.g., D-glucose & L-glucose.

Other stereoisomers have unique names, e.g., glucose, mannose, galactose, etc.



D-glucose

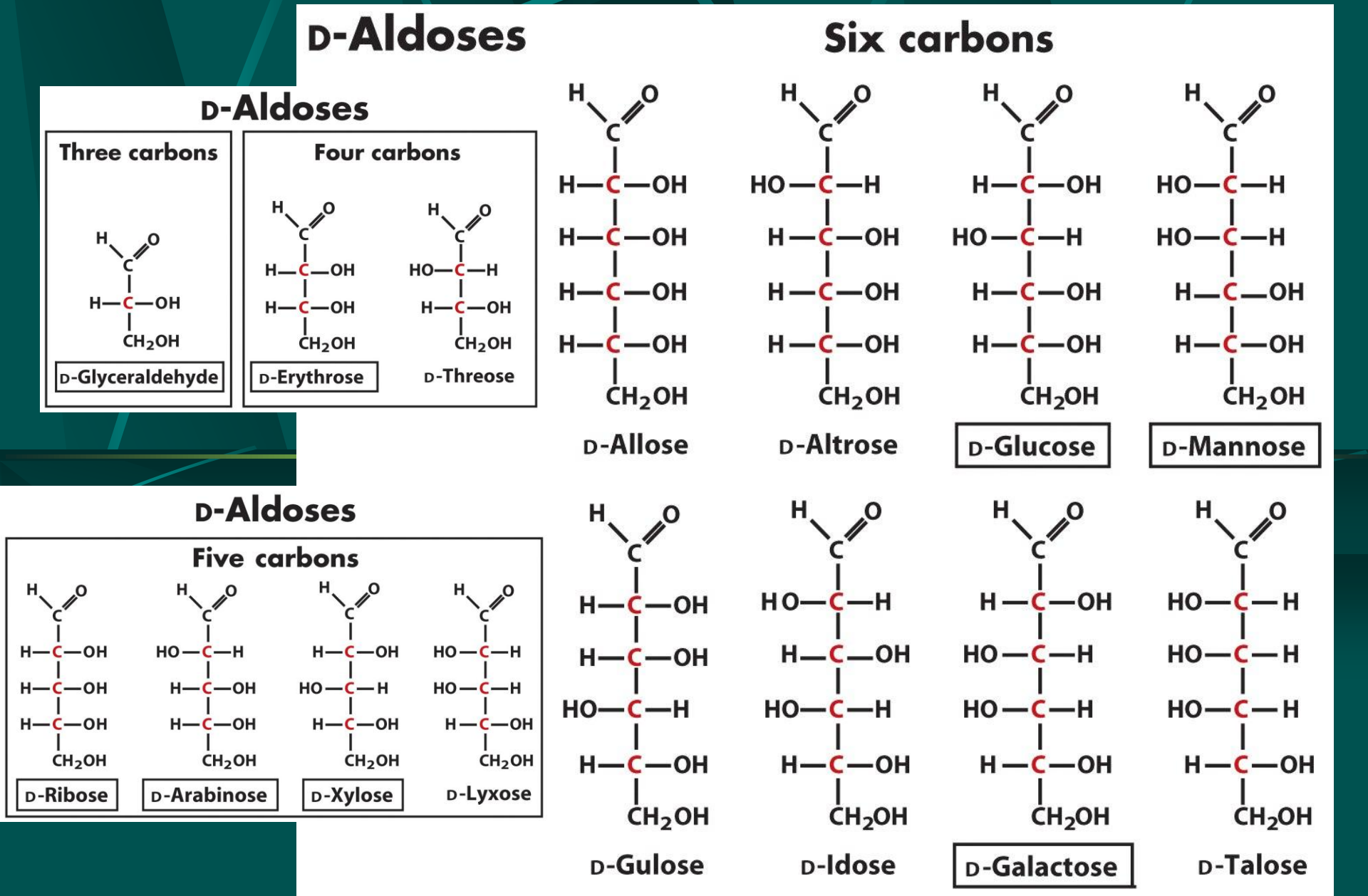


L-glucose

The number of stereoisomers is 2^n, where n is the number of asymmetric centers.

The 6-C aldoses have 4 asymmetric centers. Thus there are 16 stereoisomers (8 D-sugars and 8 L-sugars).

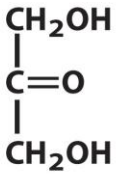
All the D-Aldoses



And All the D-Ketoses

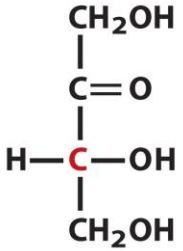
D-Ketoses

Three carbons



Dihydroxyacetone

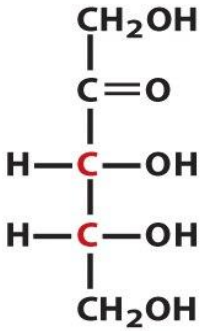
Four carbons



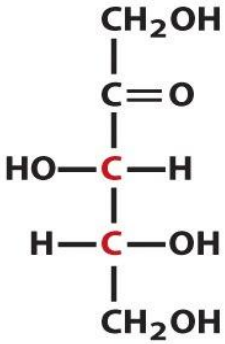
D-Erythrulose

D-Ketoses

Five carbons

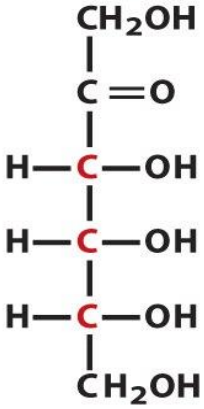


D-Ribulose

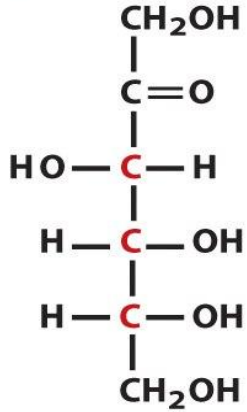


D-Xylulose

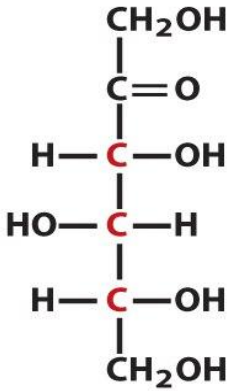
Six carbons



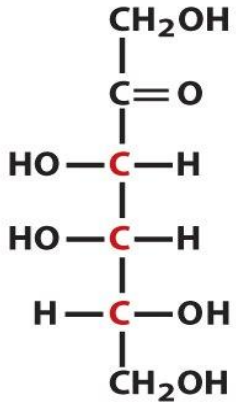
D-Psicose



D-Fructose

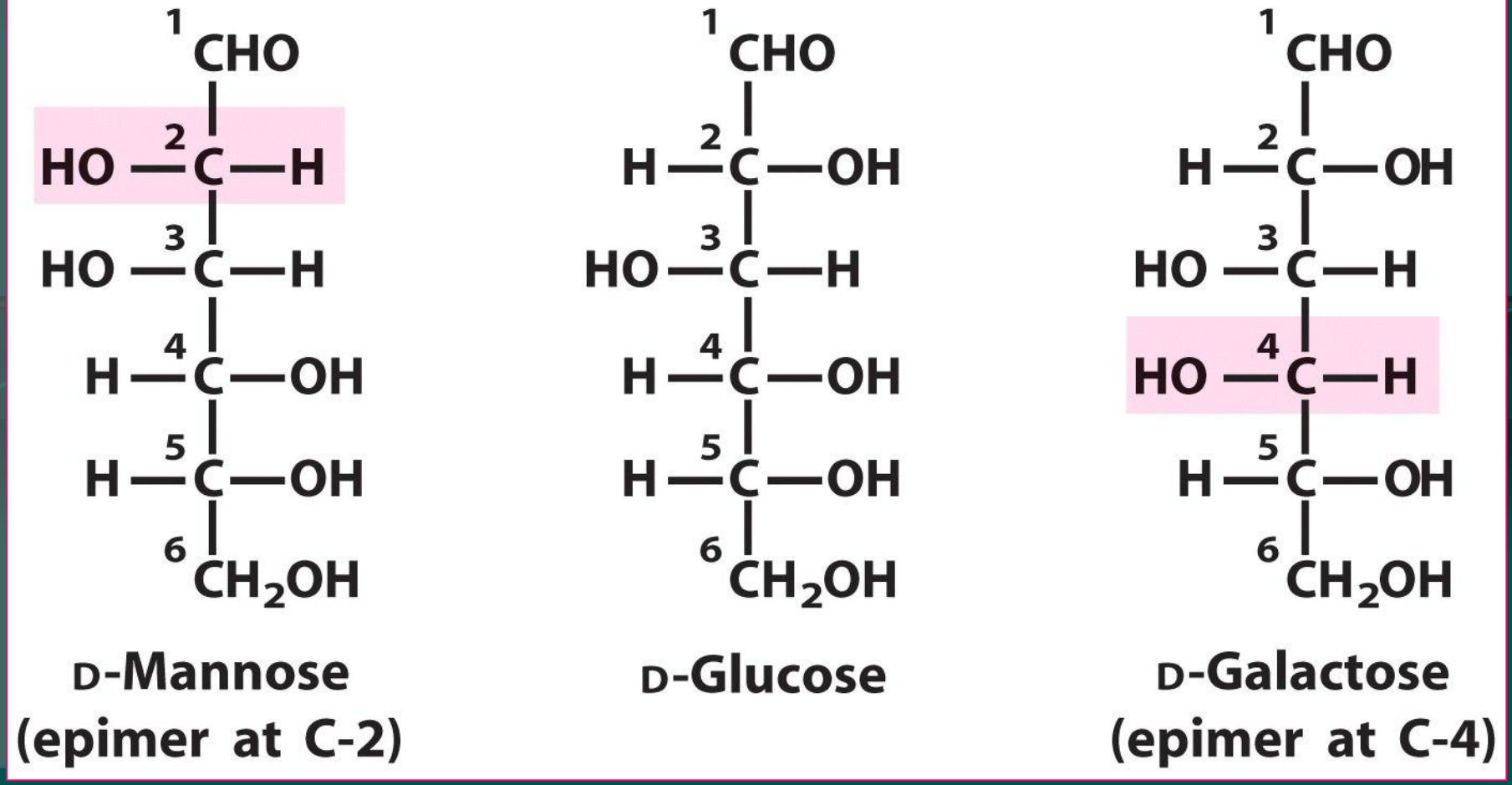


D-Sorbose



D-Tagatose

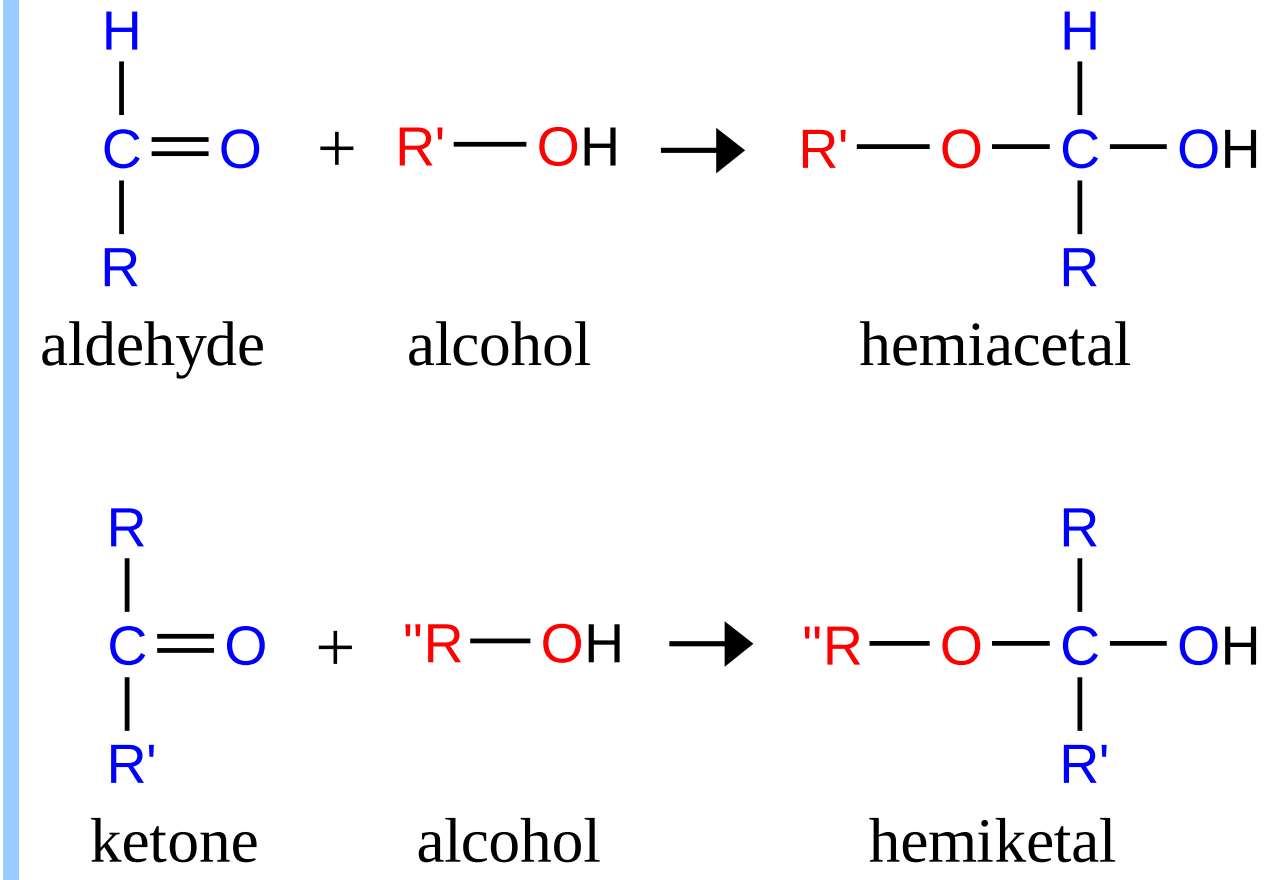
Epimers Differ at a Single Chiral Position



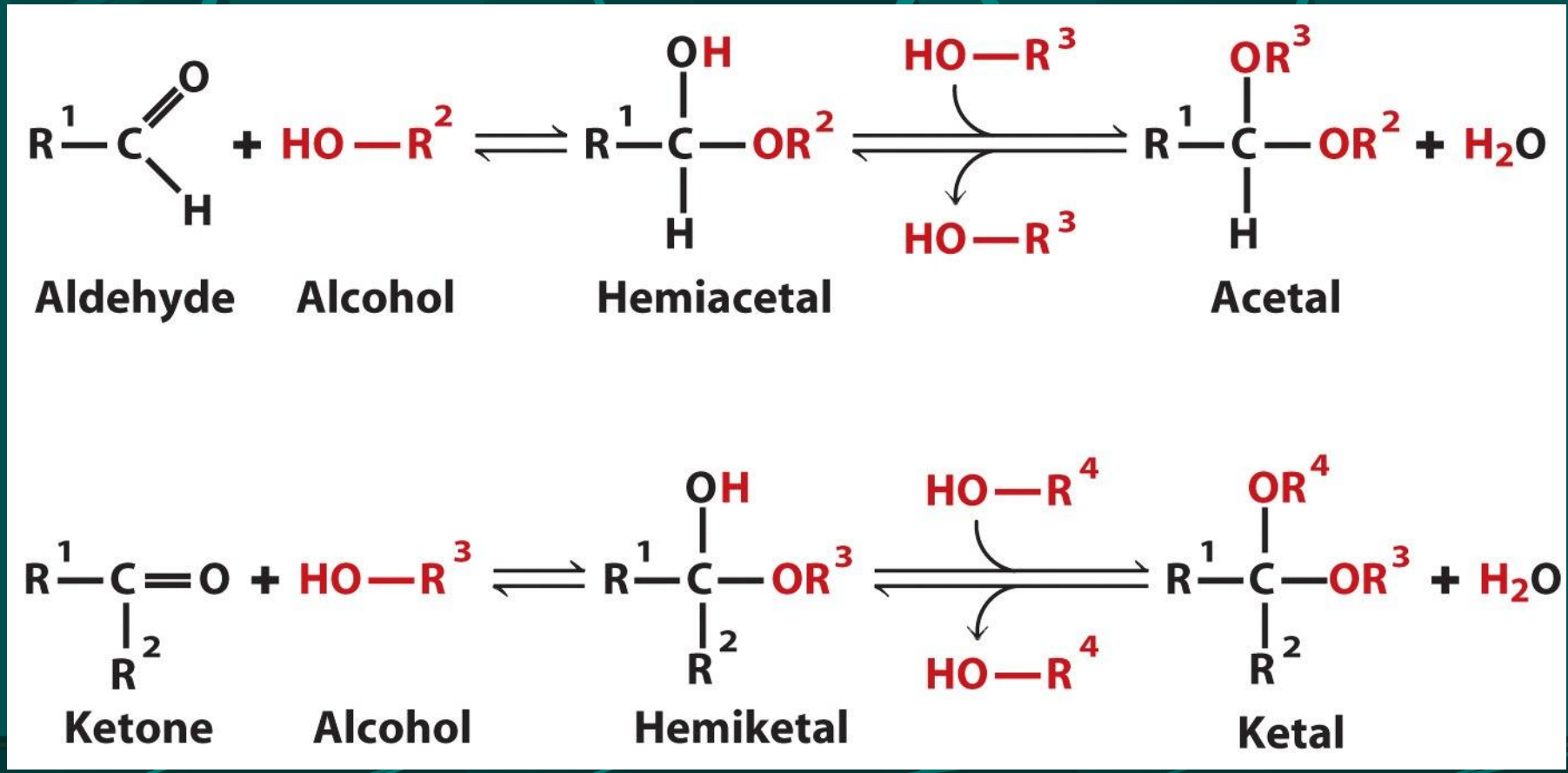
Hemiacetal & hemiketal formation

An aldehyde can react with an alcohol to form a hemiacetal.

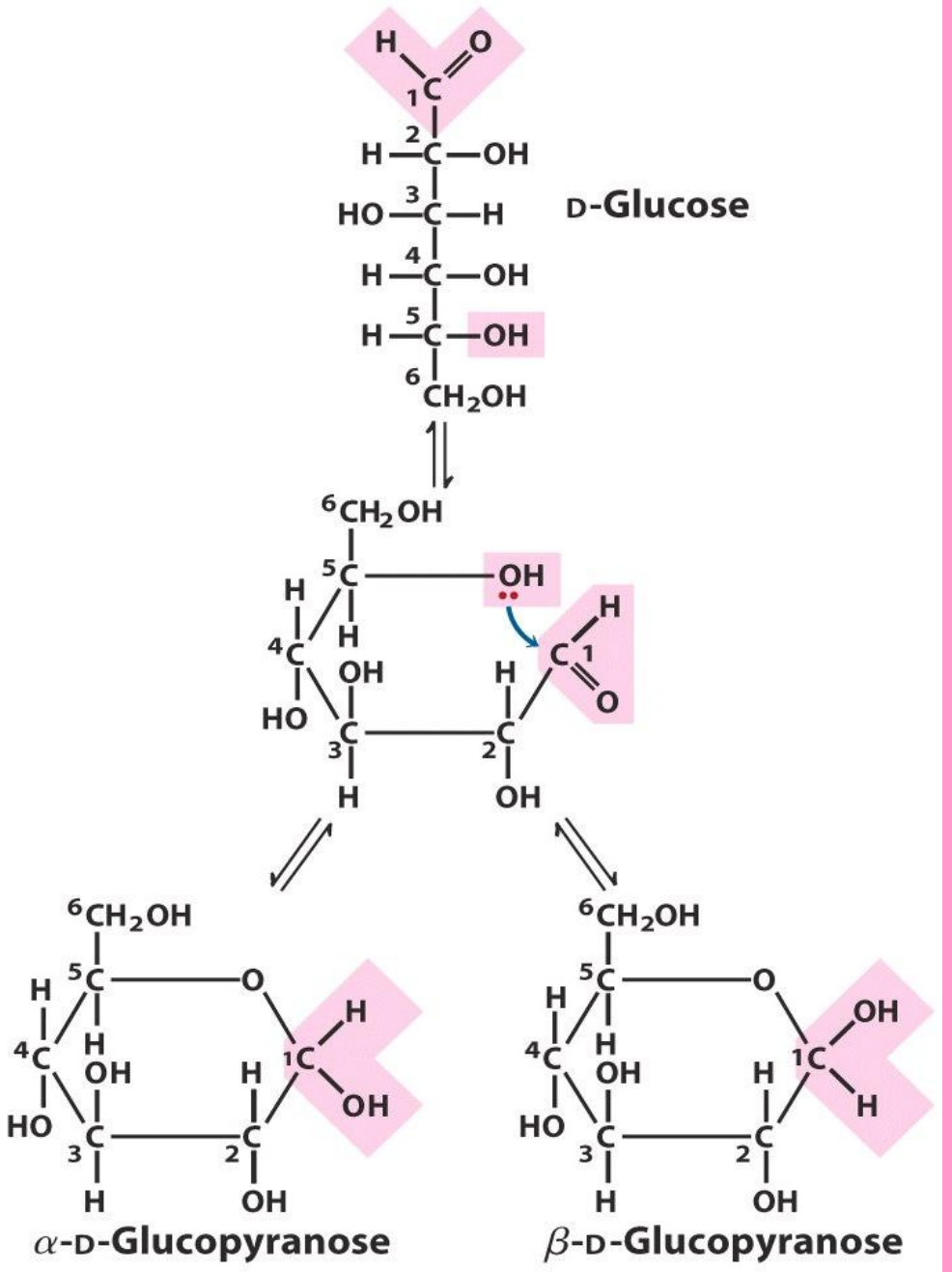
A ketone can react with an alcohol to form a hemiketal.



Formation of hemiacetals and hemiketals

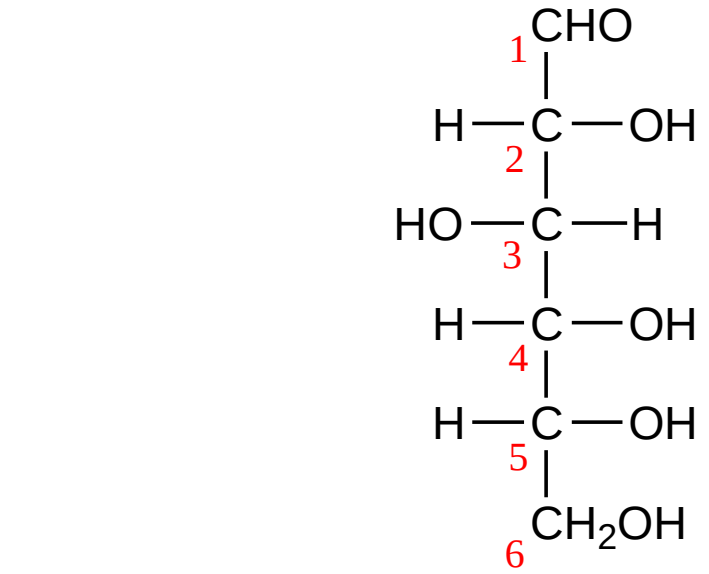


Cyclic Forms of Glucose With Hemiacetal Linkage

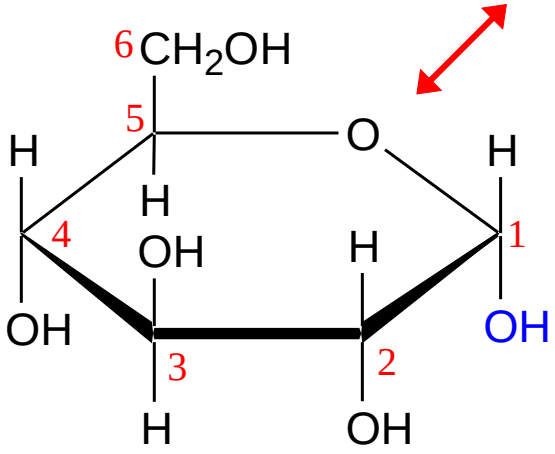


Pentoses and hexoses can cyclize as the ketone or aldehyde reacts with a distal OH.

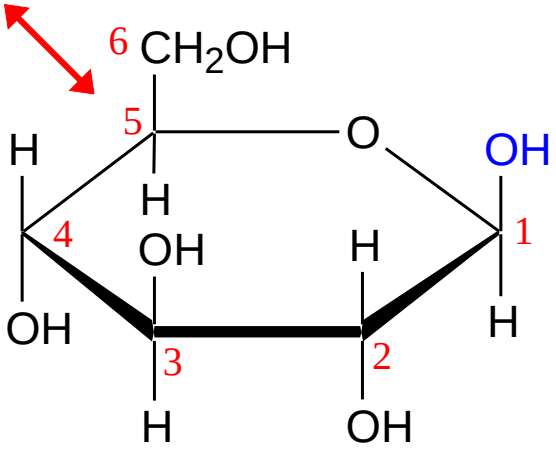
Glucose forms an intra-molecular hemiacetal, as the C1 aldehyde & C5 OH react, to form a 6-member pyranose ring, named after pyran.



D-glucose (linear form)

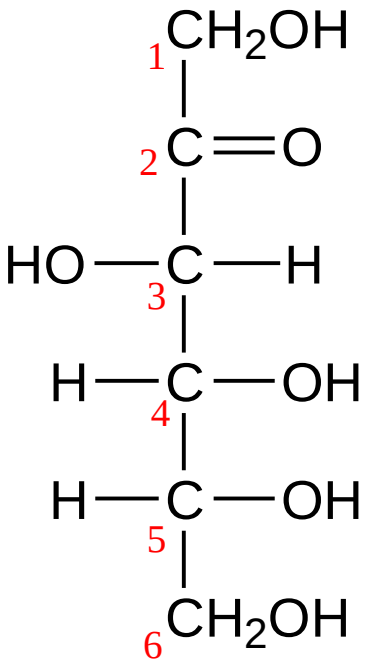


α-D-glucose

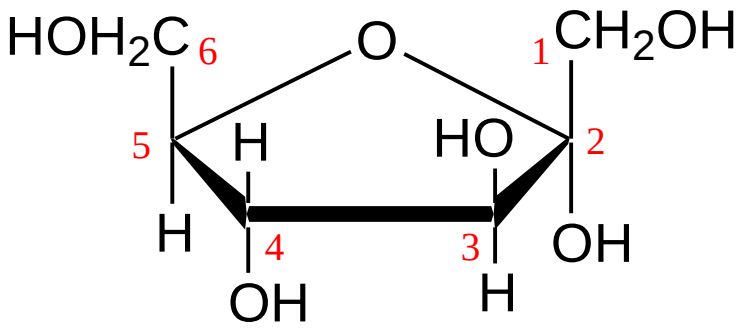


β-D-glucose

These representations of the cyclic sugars are called Haworth projections.



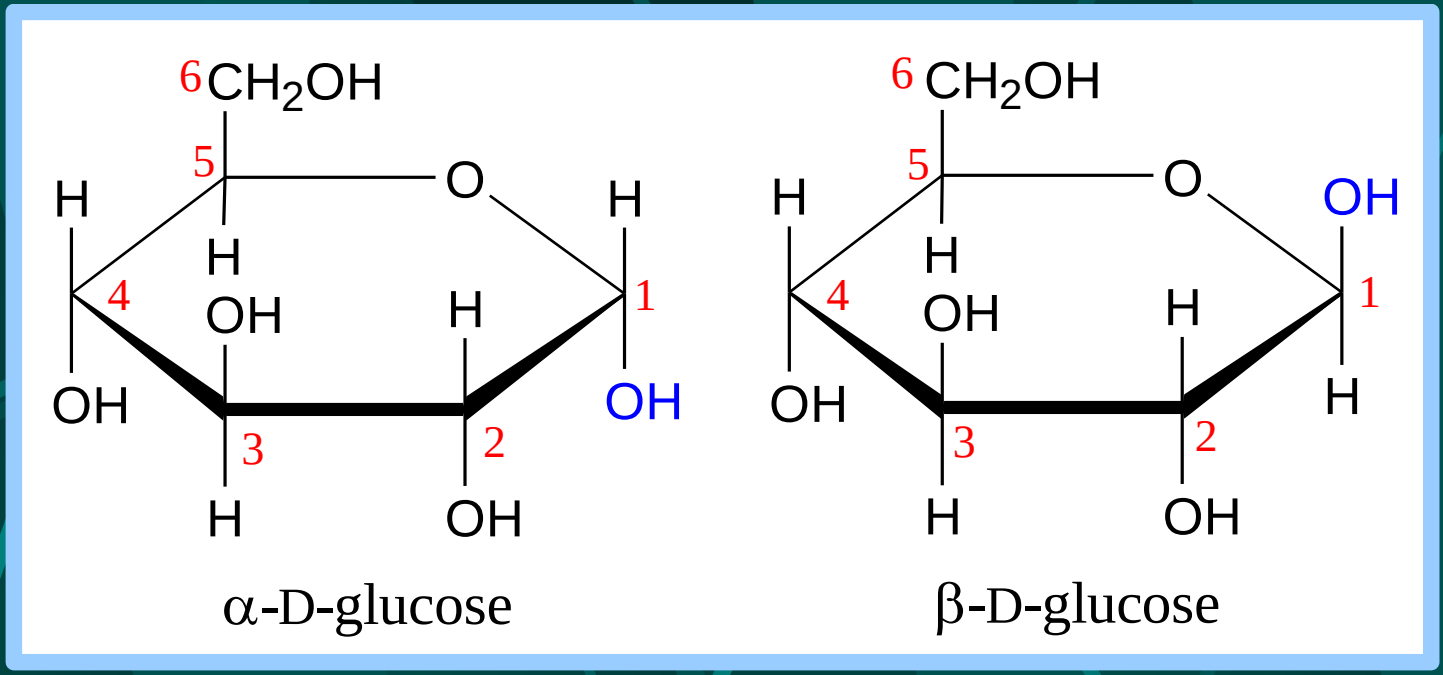
D-fructose (linear)



α-D-fructofuranose

Fructose forms either

- a 6-member pyranose ring, by reaction of the C2 keto group with the OH on C6, or
- a 5-member furanose ring, by reaction of the C2 keto group with the OH on C5.

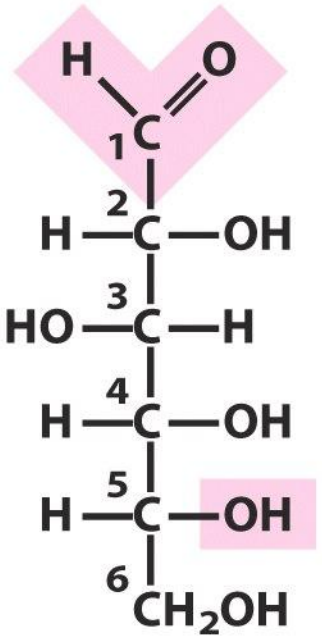


Cyclization of glucose produces a new asymmetric center at C1. The 2 stereoisomers are called anomers, a & b.

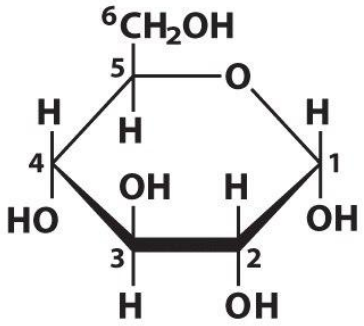
Haworth projections represent the cyclic sugars as having essentially planar rings, with the OH at the anomeric C1:

- a (OH below the ring)
- b (OH above the ring).

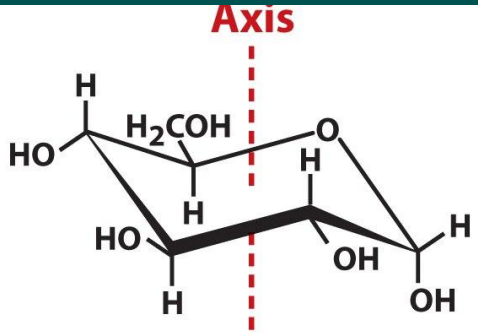
Converting Between Structural Representations



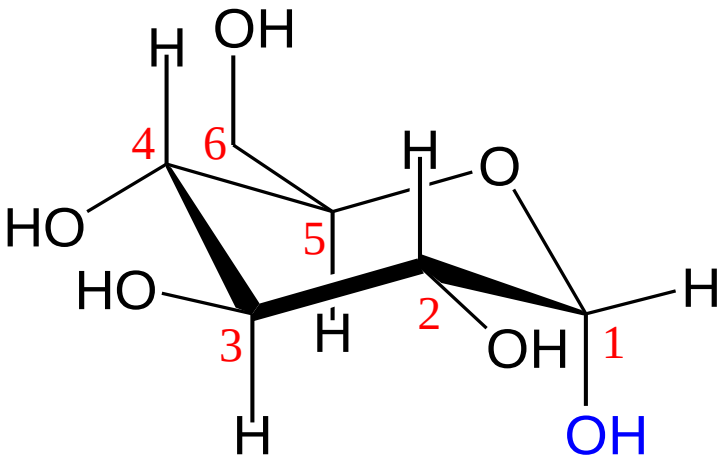
D-Glucose



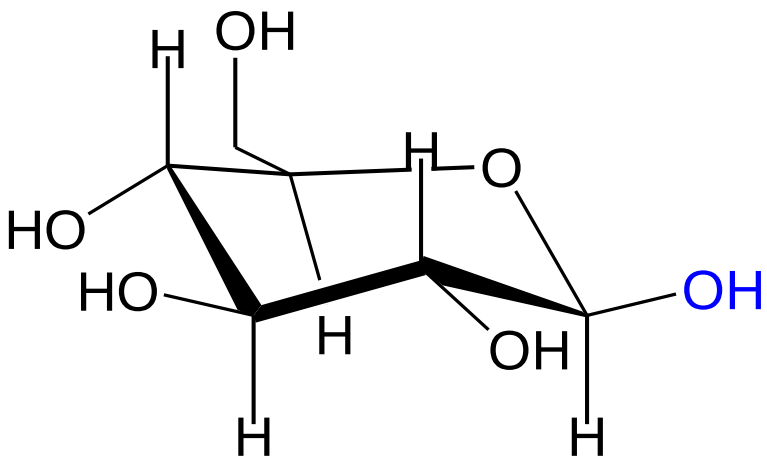
α-D-Glucopyranose



α-D-Glucopyranose



α-D-glucopyranose

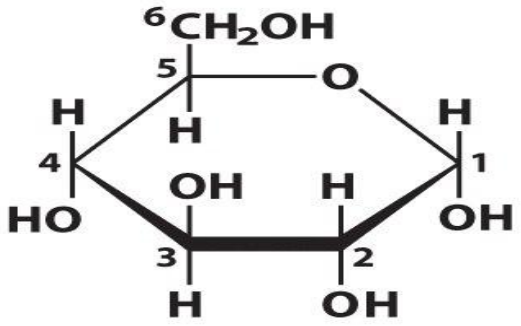


β-D-glucopyranose

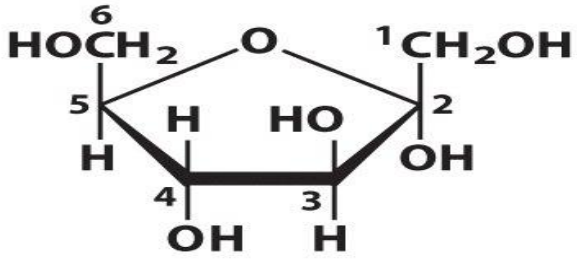
Because of the tetrahedral nature of carbon bonds, pyranose sugars actually assume a "chair" or "boat" configuration, depending on the sugar.

The representation above reflects the chair configuration of the glucopyranose ring more accurately than the Haworth projection.

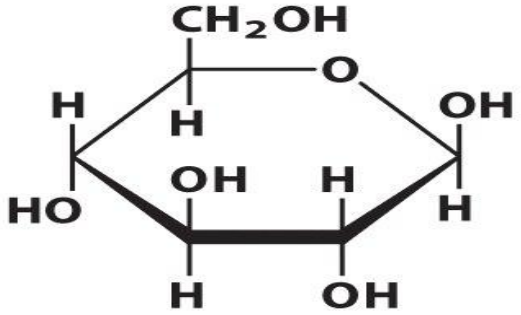
Pyranoses and Furanoses



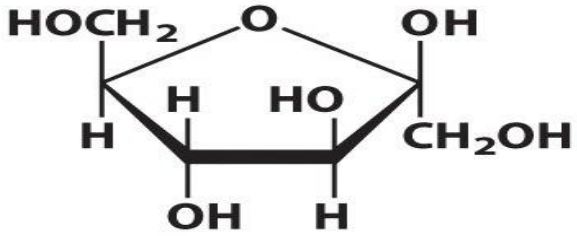
α-D-Glucopyranose



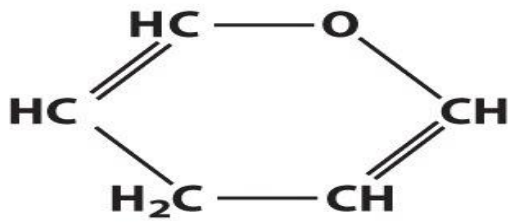
α-D-Fructofuranose



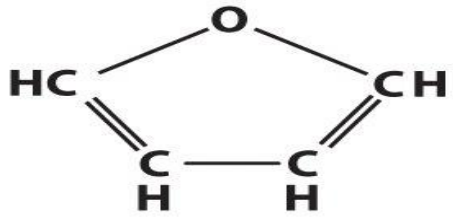
β-D-Glucopyranose



β-D-Fructofuranose

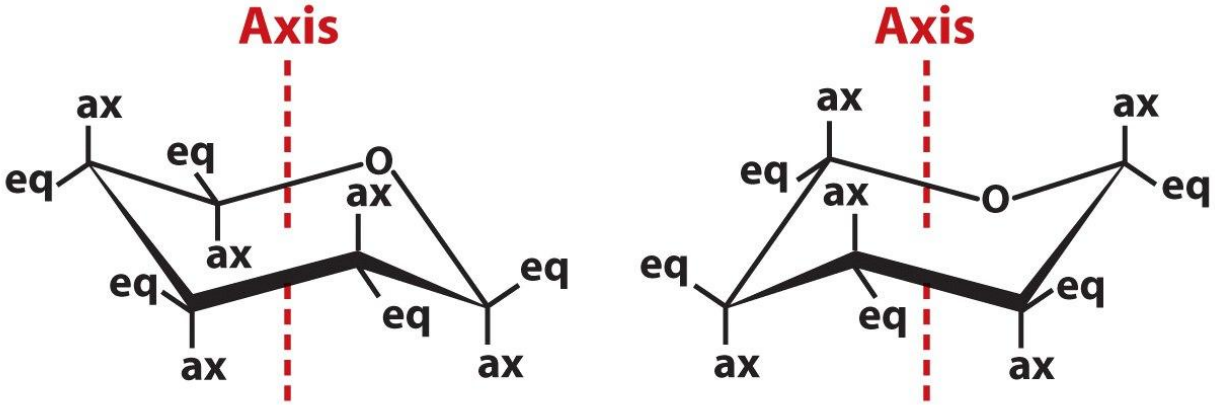


Pyran

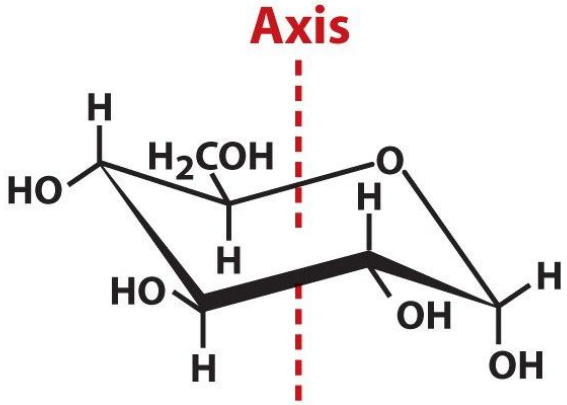


Furan

Two possible chair conformations

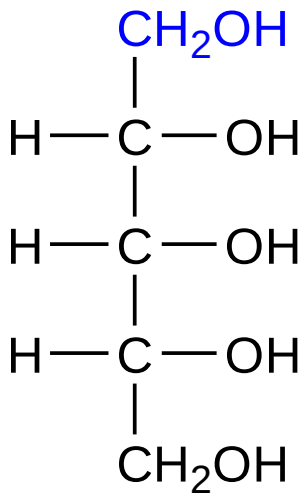


Two possible chair forms

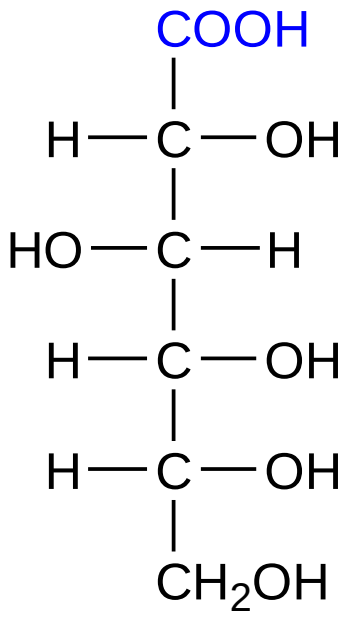


α-D-Glucopyranose

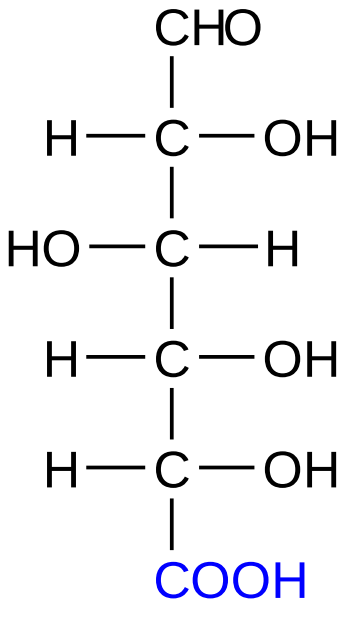
Sugar derivatives



D-ribitol



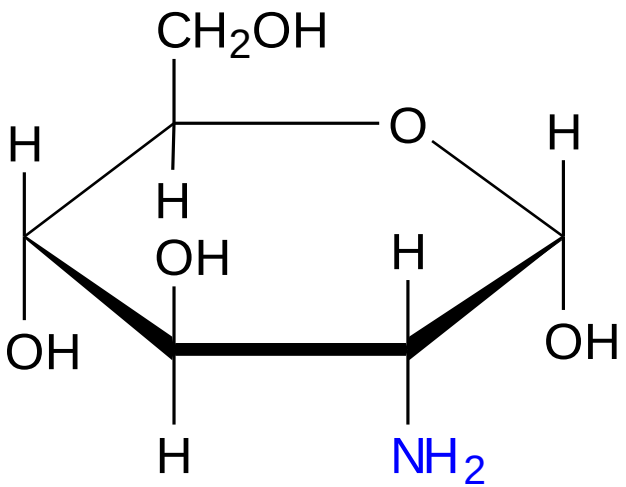
D-gluconic acid



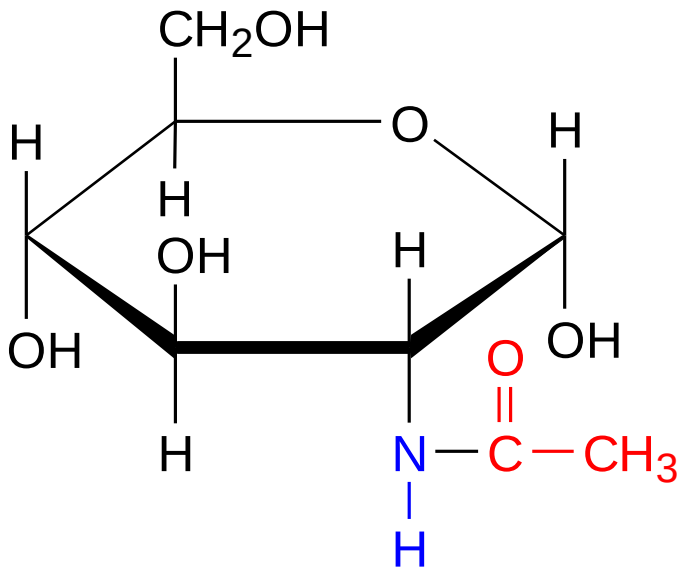
D-glucuronic acid

- sugar alcohol - lacks an aldehyde or ketone; e.g., ribitol.
- sugar acid - the aldehyde at C1, or OH at C6, is oxidized to a carboxylic acid; e.g., gluconic acid, glucuronic acid.

Sugar derivatives



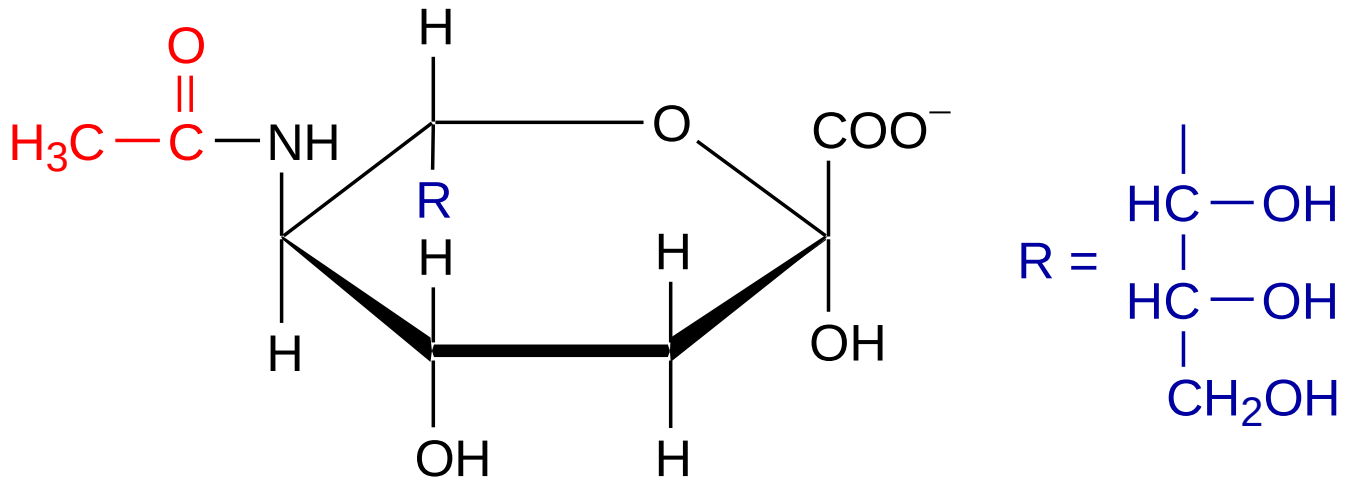
α-D-glucosamine



α-D-*N*-acetylglucosamine

amino sugar - an amino group substitutes for a hydroxyl. An example is glucosamine.

The amino group may be acetylated, as in *N*-acetylglucosamine.

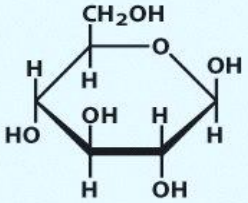


N-acetylneuraminate (sialic acid)

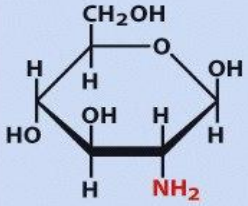
N-acetylneuraminate (N-acetylneuraminic acid, also called sialic acid) is often found as a terminal residue of oligosaccharide chains of glycoproteins.

Sialic acid imparts negative charge to glycoproteins, because its carboxyl group tends to dissociate a proton at physiological pH, as shown here.

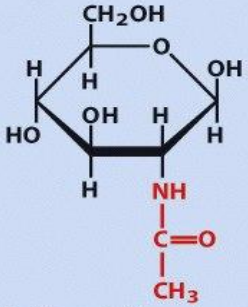
Glucose family



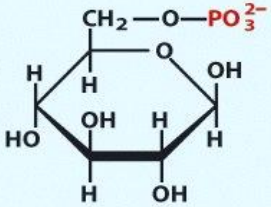
β-D-Glucose



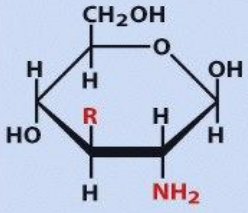
β-D-Glucosamine



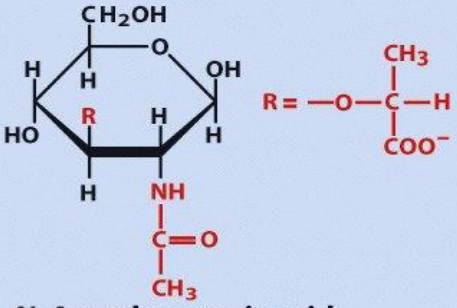
N-Acetyl-β-D-glucosamine



β-D-Glucose 6-phosphate

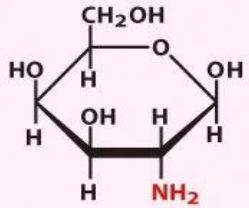


Muramic acid

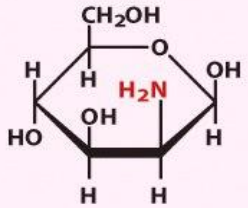


N-Acetylmuramic acid

Amino sugars

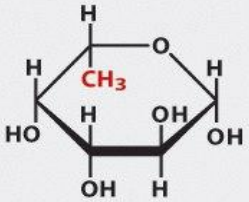


β-D-Galactosamine

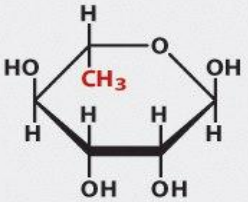


β-D-Mannosamine

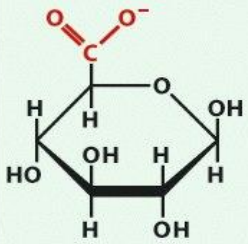
Deoxy sugars



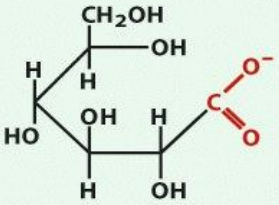
β-L-Fucose



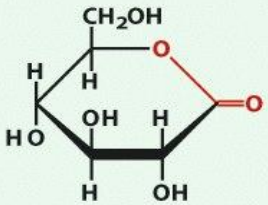
α-L-Rhamnose



β-D-Glucuronate

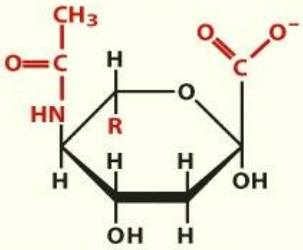


D-Gluconate



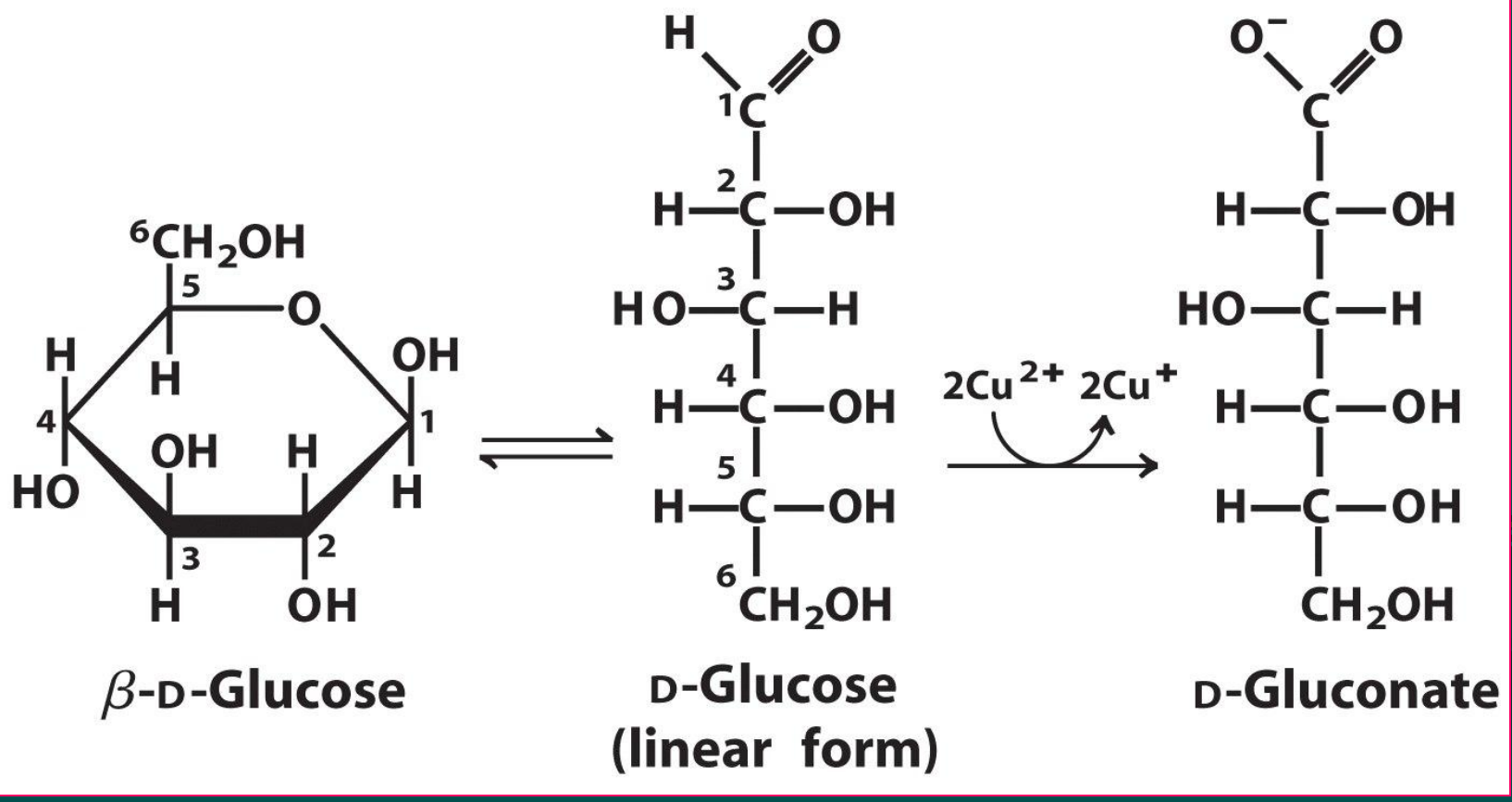
D-Glucono-δ-lactone

Acidic sugars



N-Acetylneuraminic acid (a sialic acid)

Sugars as Reducing Agents



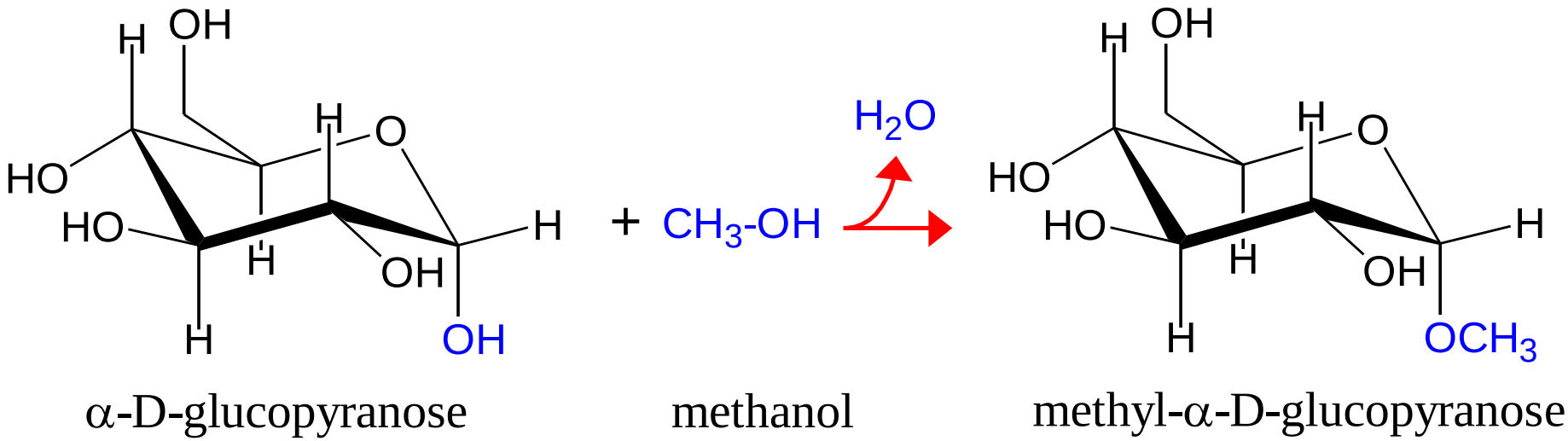
$$\text{D-Glucose} + O_2 \xrightarrow{\text{glucose oxidase}} \text{D-Glucono-}\delta\text{-lactone} + H_2O_2$$

Glycosidic Bonds

The anomeric hydroxyl and a hydroxyl of another sugar or some other compound can join together, splitting out water to form a glycosidic bond:

R-OH + HO-R' → R-O-R' + H_2O

E.g., methanol reacts with the anomeric OH on glucose to form methyl glucoside (methyl-glucopyranose).

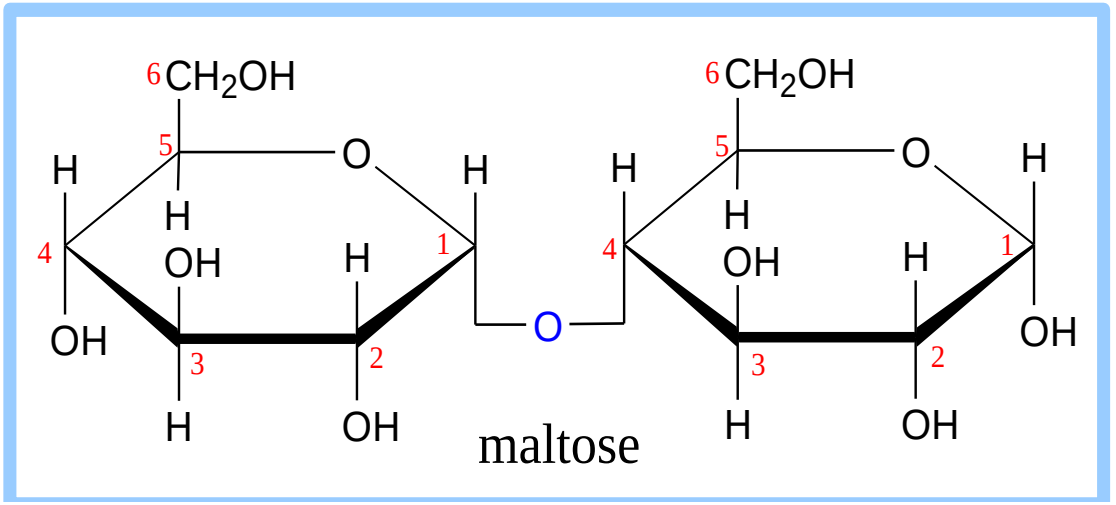


α-D-glucopyranose methanol methyl-α-D-glucopyranose

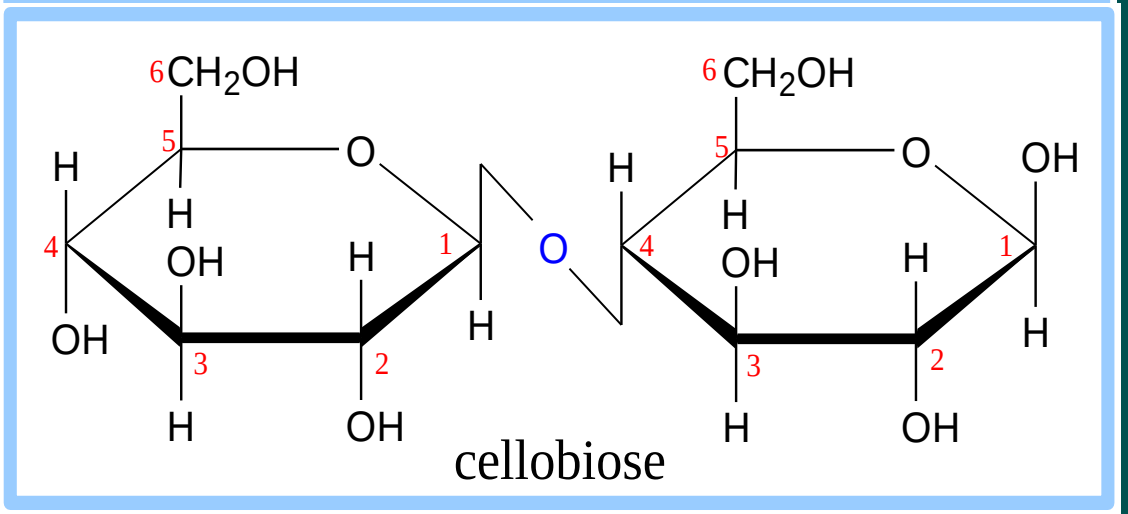
Disaccharides:

Maltose, a cleavage product of starch (e.g., amylose), is a disaccharide with an α(1→ 4) glycosidic link between C1 - C4 OH of 2 glucoses.

It is the **α** anomer (C1 O points down).



maltose



cellobiose

Cellobiose, a product of cellulose breakdown, is the otherwise equivalent b anomer (O on C1 points up).

The b(1® 4) glycosidic linkage is represented as a zig-zag, but one glucose is actually flipped over relative to the other.

Other disaccharides include:

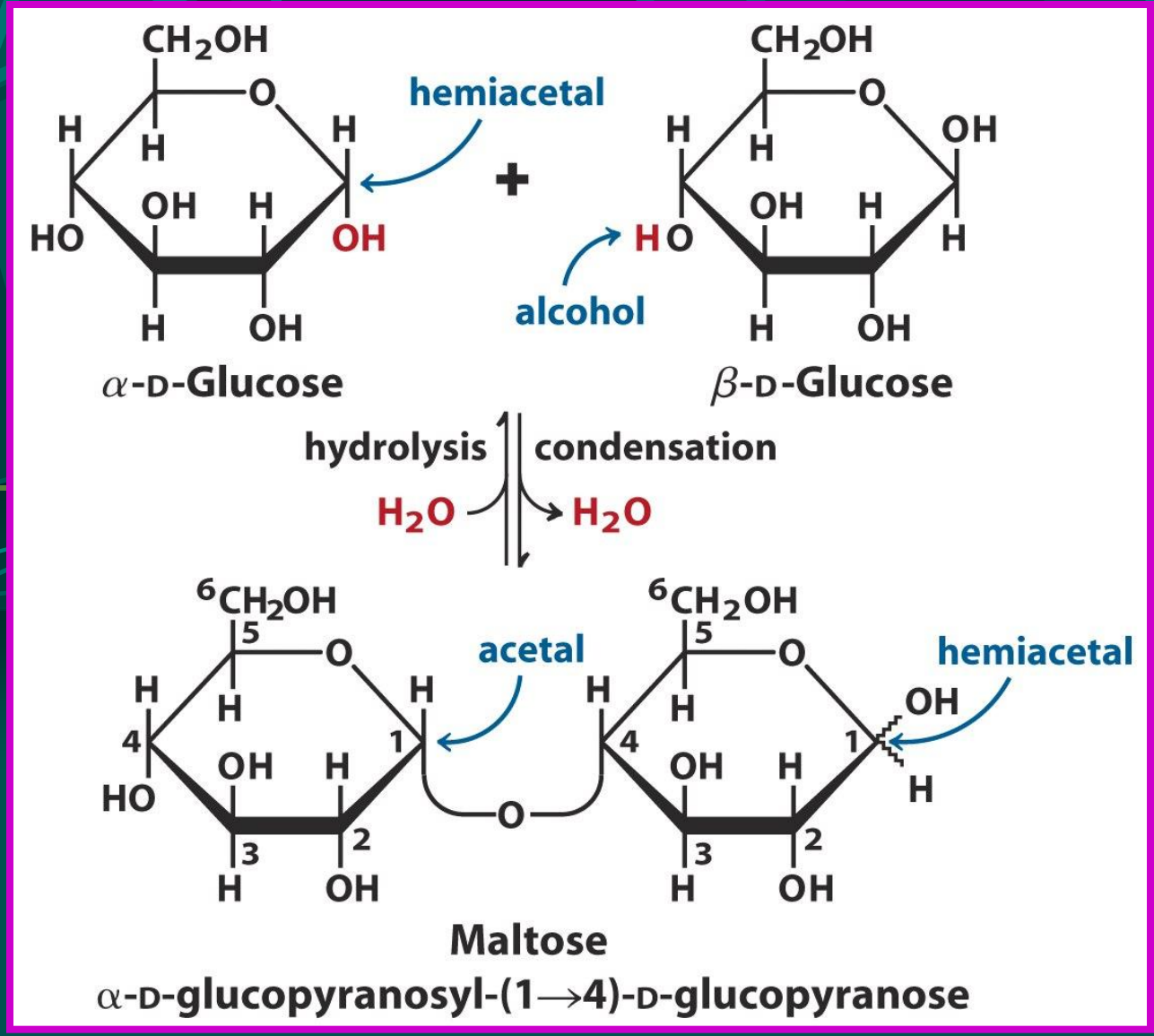
- **Sucrose**, common table sugar, has a glycosidic bond linking the anomeric hydroxyls of glucose & fructose.

  Because the configuration at the anomeric C of glucose is a (O points down from ring), the linkage is a(1→2).

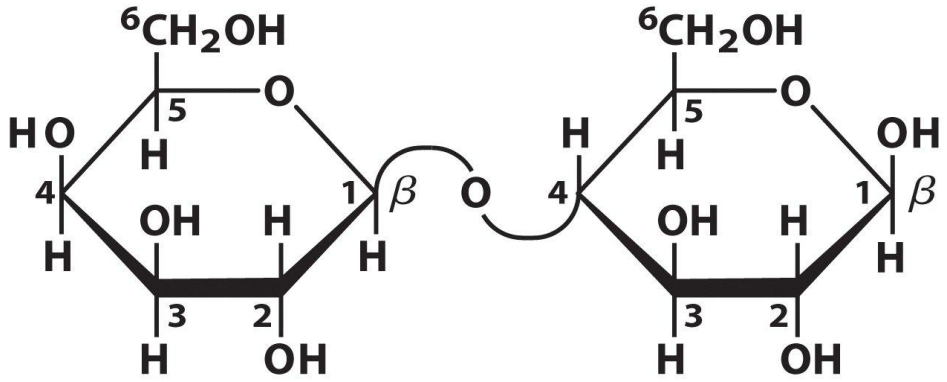
  The full name of sucrose is a-D-glucopyranosyl-(1→2)-b-D-fructopyranose.)

- **Lactose**, milk sugar, is composed of galactose & glucose, with b(1→4) linkage from the anomeric OH of galactose. Its full name is b-D-galactopyranosyl-(1→ 4)-a-D-glucopyranose

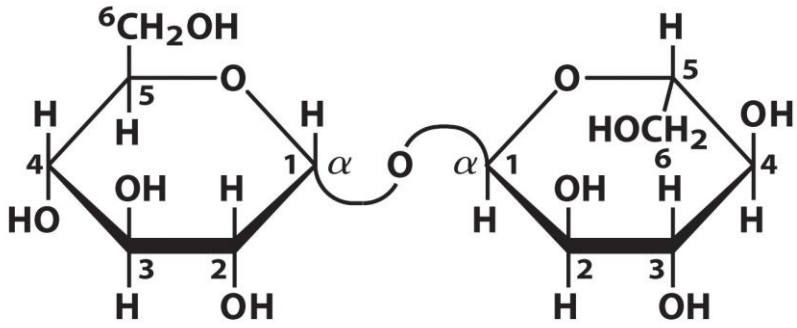
Two Molecules of Glucose Form Maltose



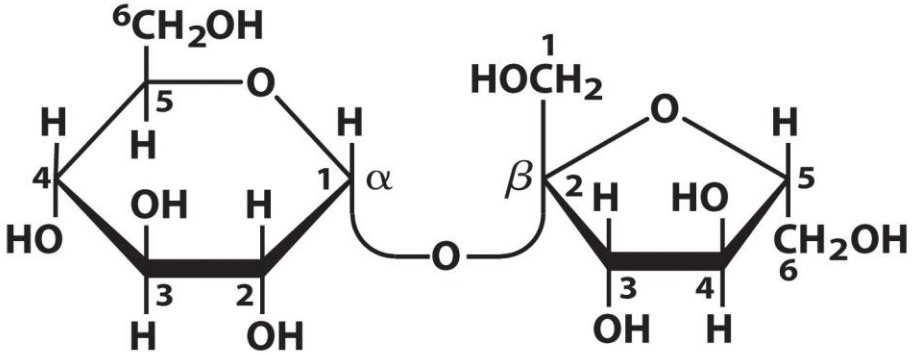
Common Disaccharides



Lactose (β form)
β-D-galactopyranosyl-(1$\rightarrow$4)-β-D-glucopyranose
Gal(β1$\rightarrow$4)Glc

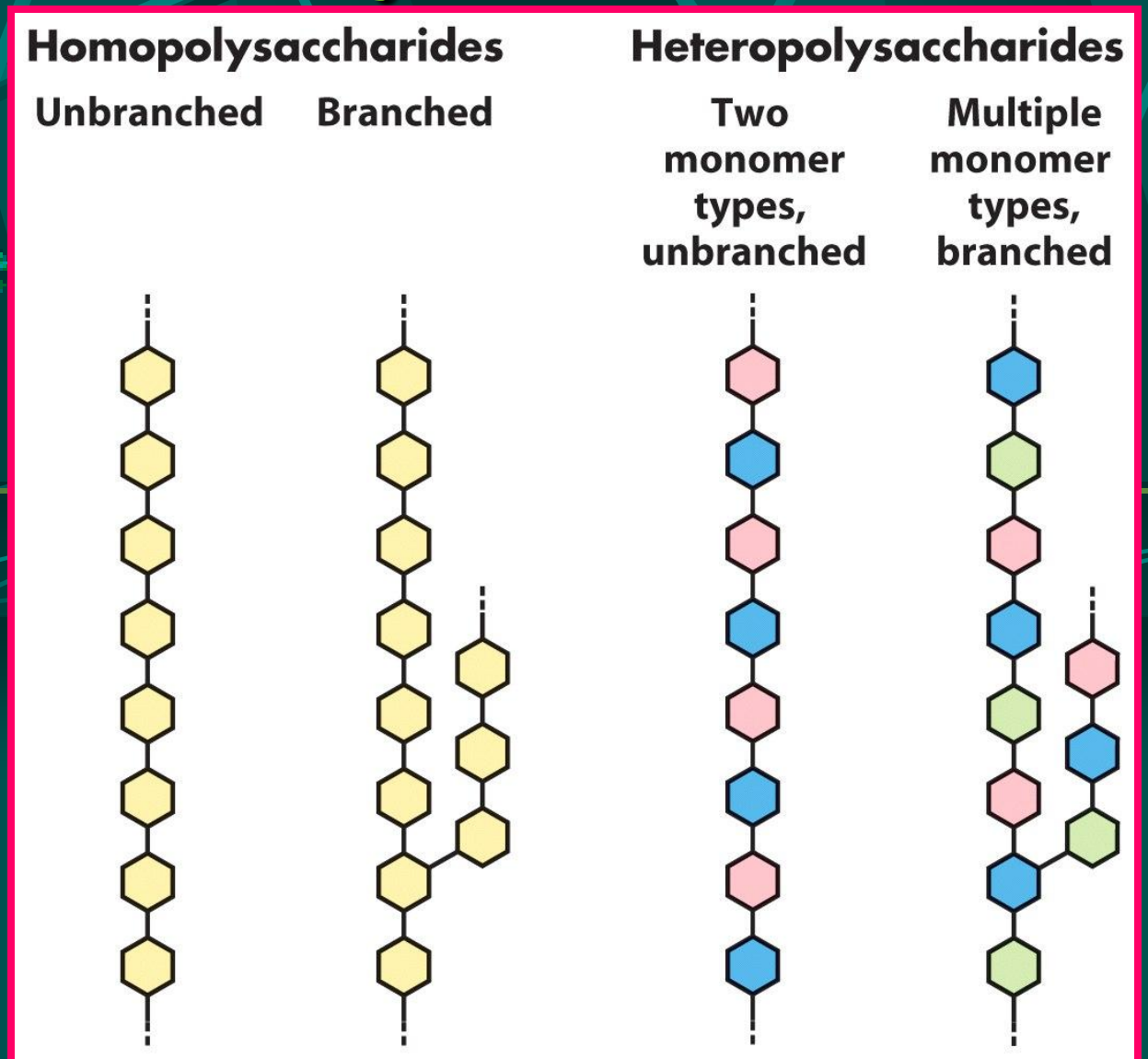


Trehalose
α-D-glucopyranosyl α-D-glucopyranoside
Glc(α1$\leftrightarrow$1α)Glc

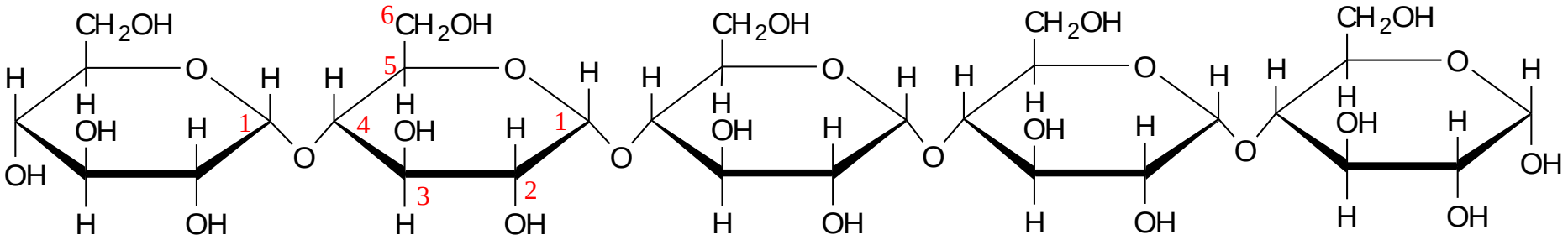


Sucrose
α-D-glucopyranosyl β-D-fructofuranoside
Glc(α1$\leftrightarrow$2β)Fru

Two General Types of Polysaccharides



Polysaccharides

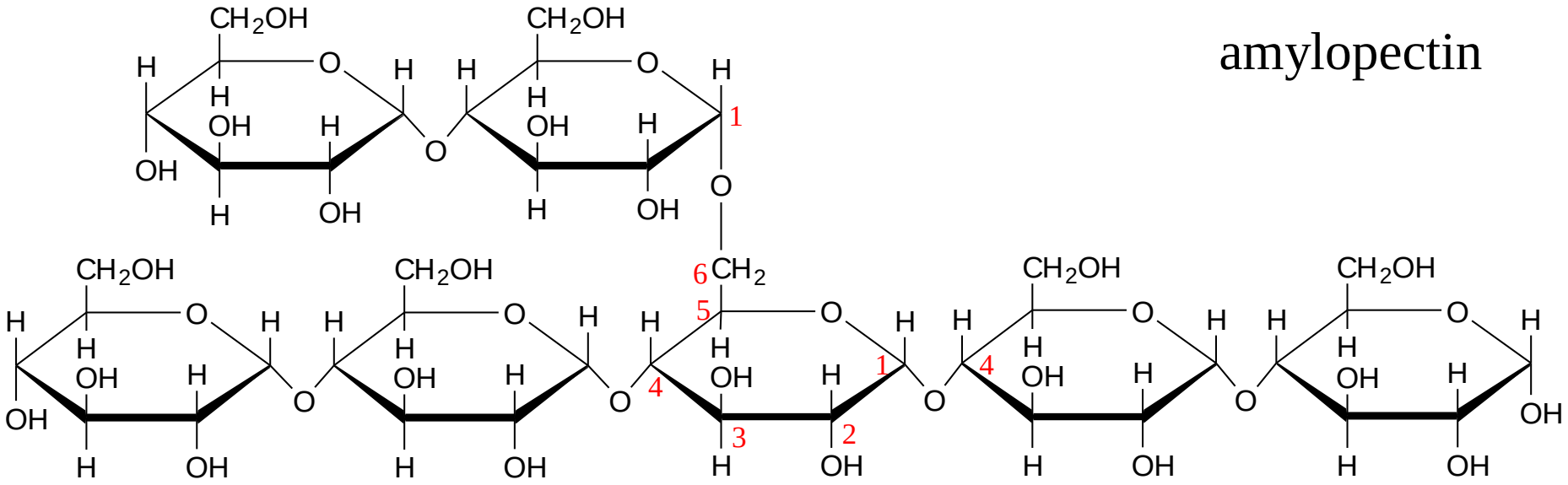


amylose

Plants store glucose as amylose or amylopectin, glucose polymers collectively called starch. Glucose storage in polymeric form minimizes osmotic effects.

Amylose is a glucose polymer with a(1→4) linkages. It adopts a helical conformation.

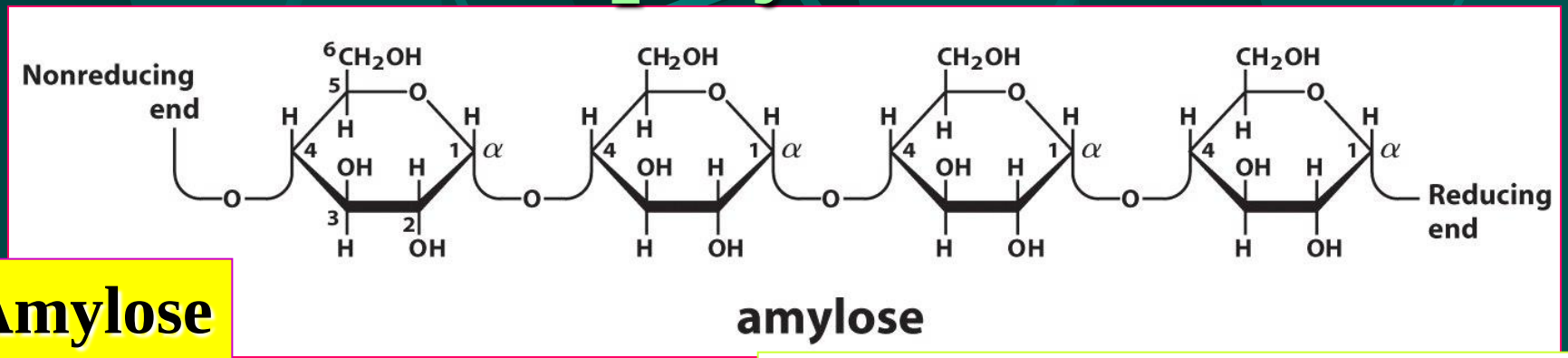
The end of the polysaccharide with an anomeric C1 not involved in a glycosidic bond is called the reducing end.



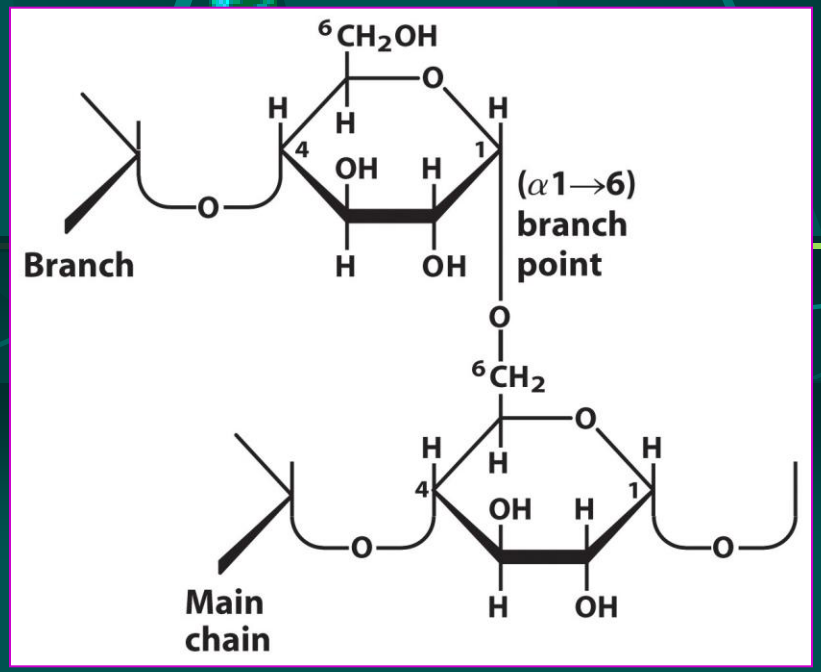
Amylopectin is a glucose polymer with mainly a(1→4) linkages, but it also has branches formed by a(1→6) linkages. Branches are generally longer than shown above.

The branches produce a compact structure & provide multiple chain ends at which enzymatic cleavage can occur.

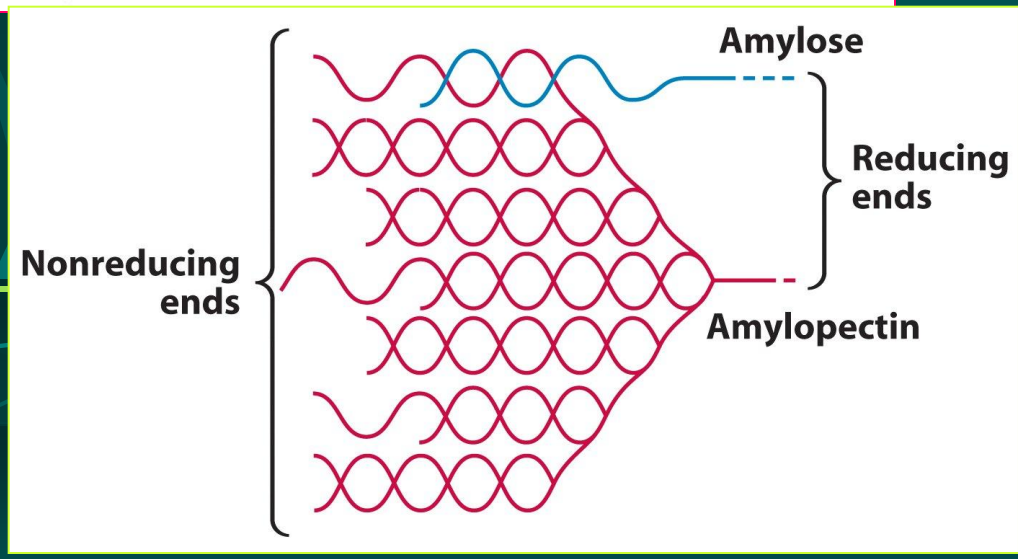
Starch: A polymer of Glucose



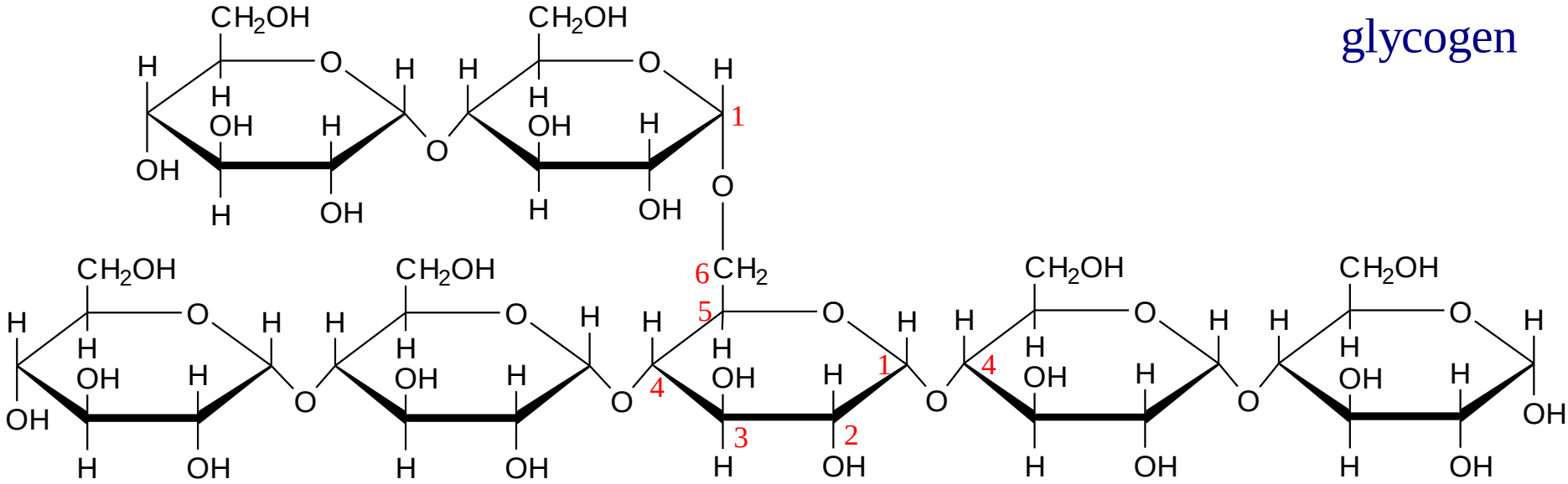
Amylose



Amylopectin



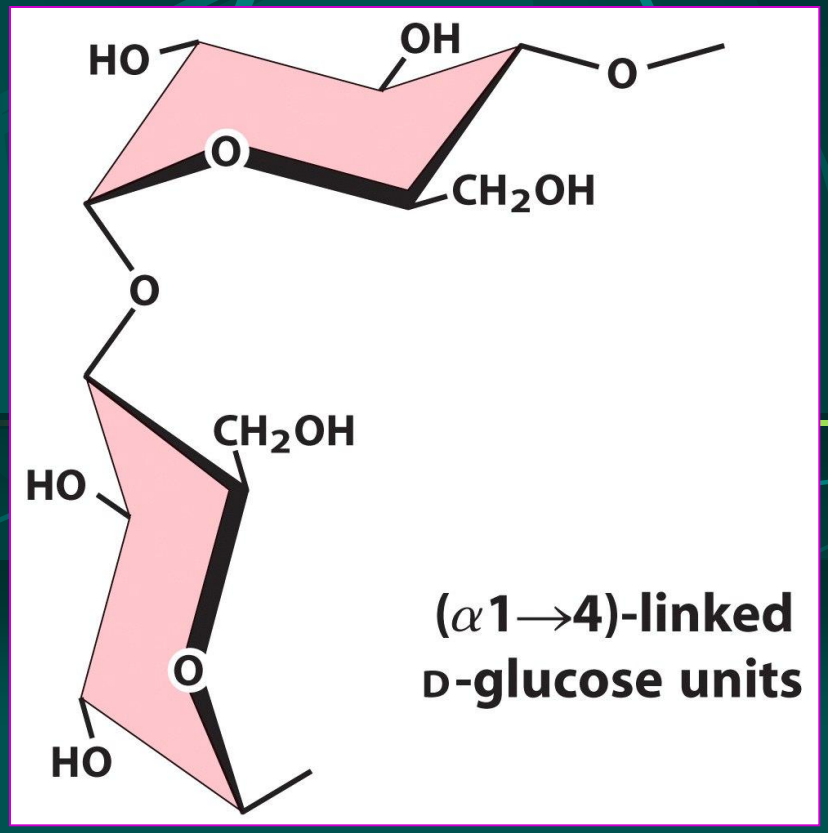
Glycogen is also branched chains of Glc(a1->4)Glc

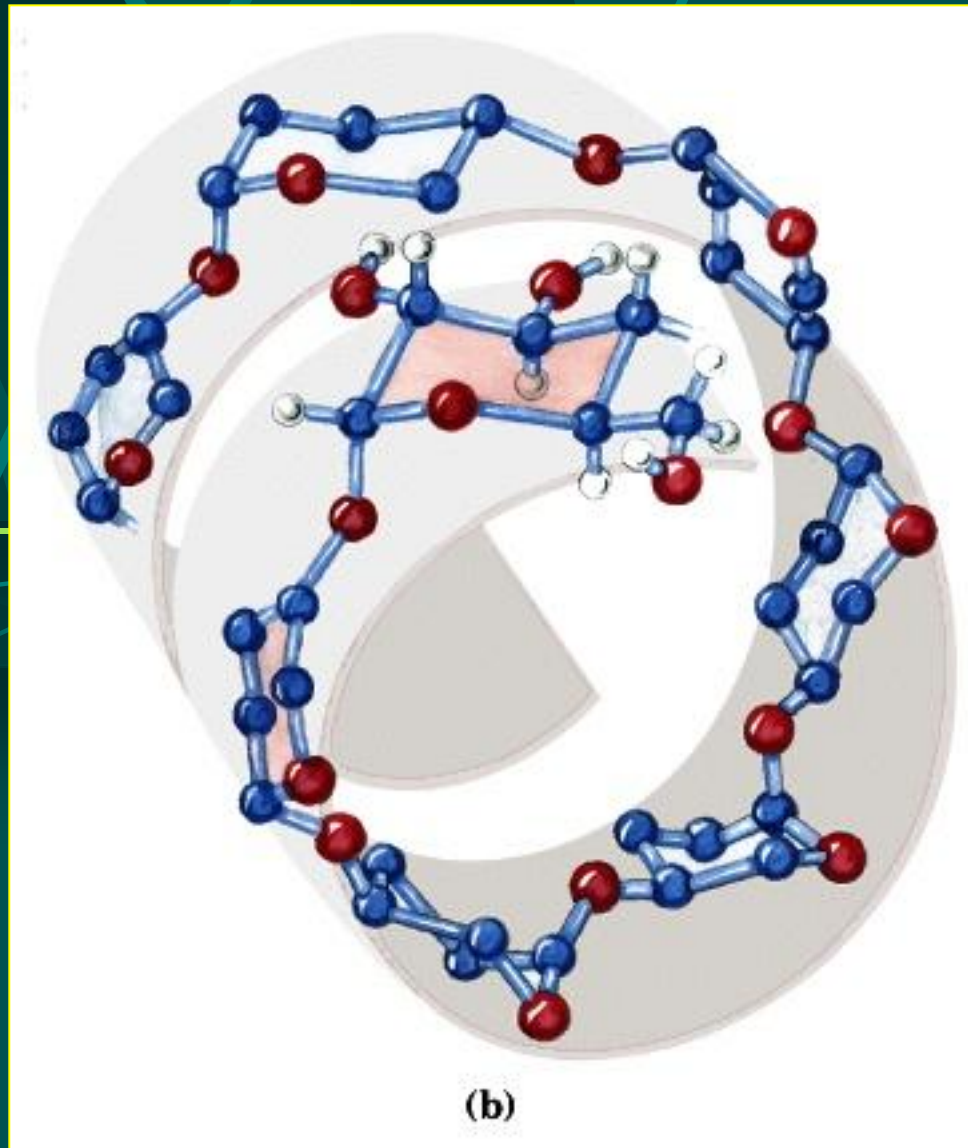


Glycogen, the glucose storage polymer in animals, is similar in structure to amylopectin. But glycogen has more a(1→6) branches.

The highly branched structure permits rapid release of glucose from glycogen stores, e.g., in muscle during exercise. The ability to rapidly mobilize glucose is more essential to animals than to plants.

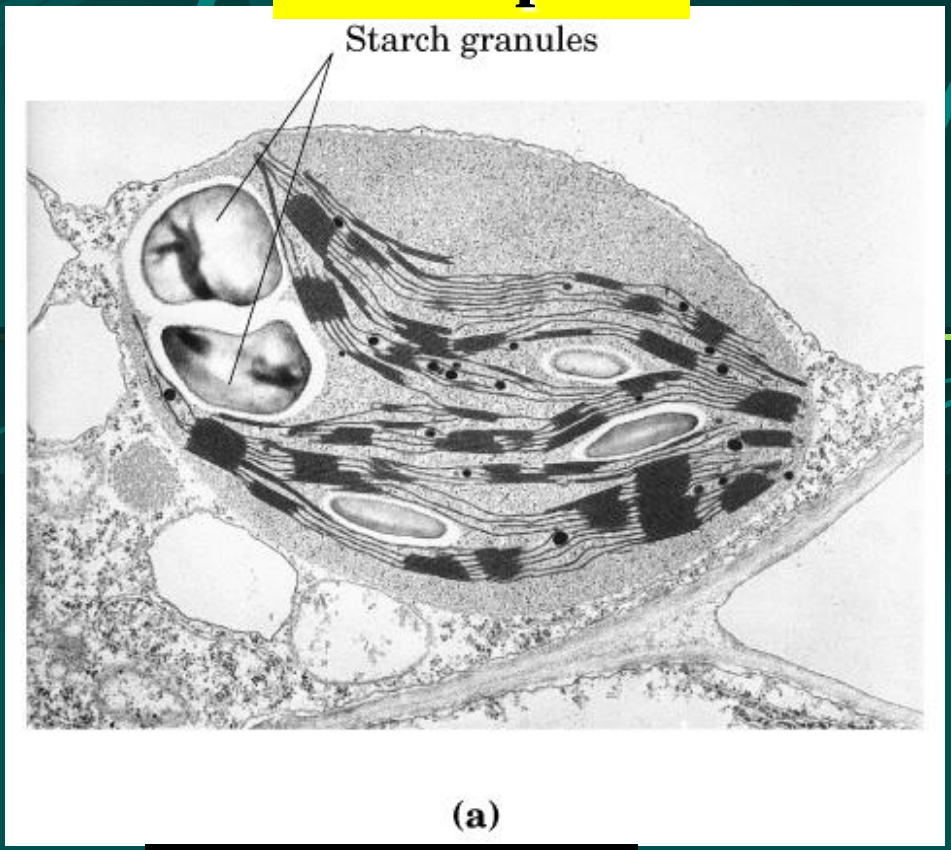
Starch and Glycogen Form Helical Structures



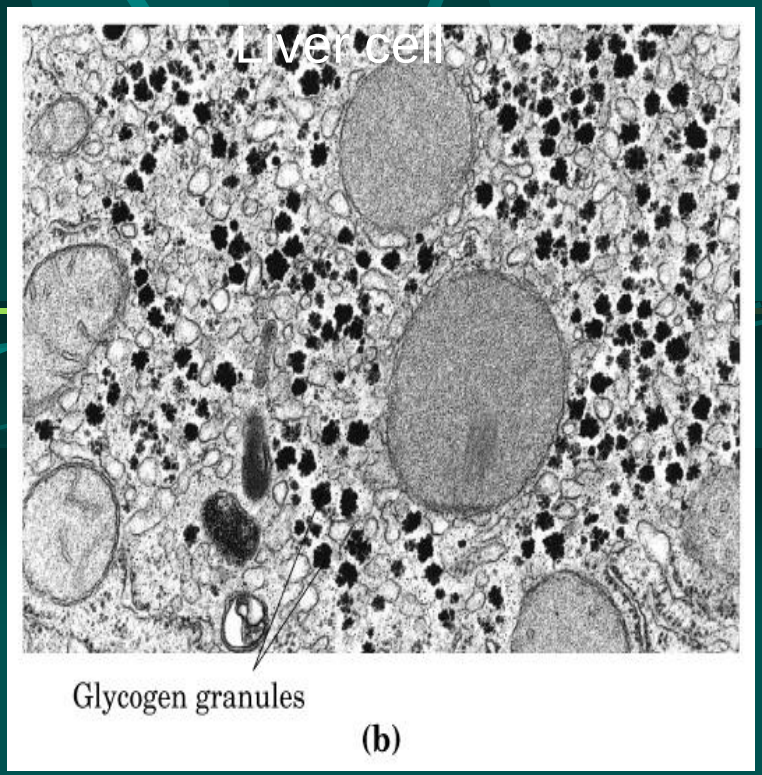


Starch and Glycogen Both Are Stored in Granules

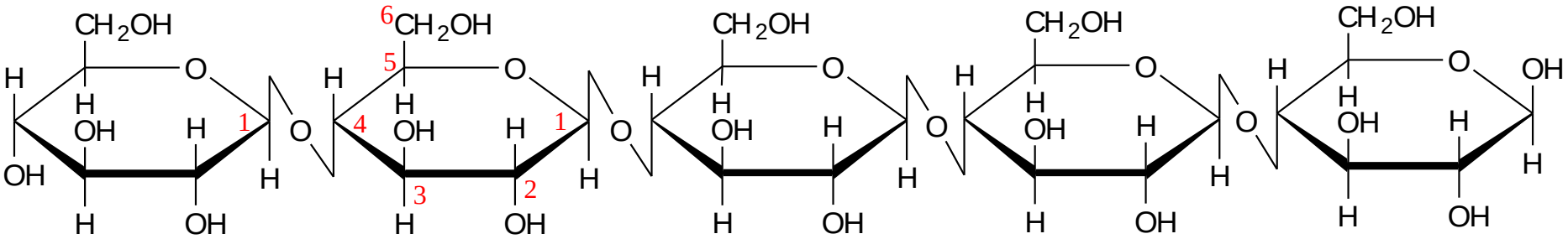
Chloroplast



1 µm granules

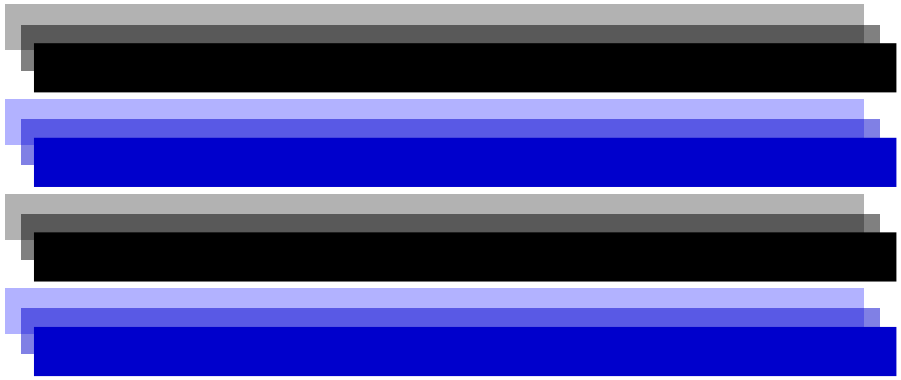


0.1 µm granules

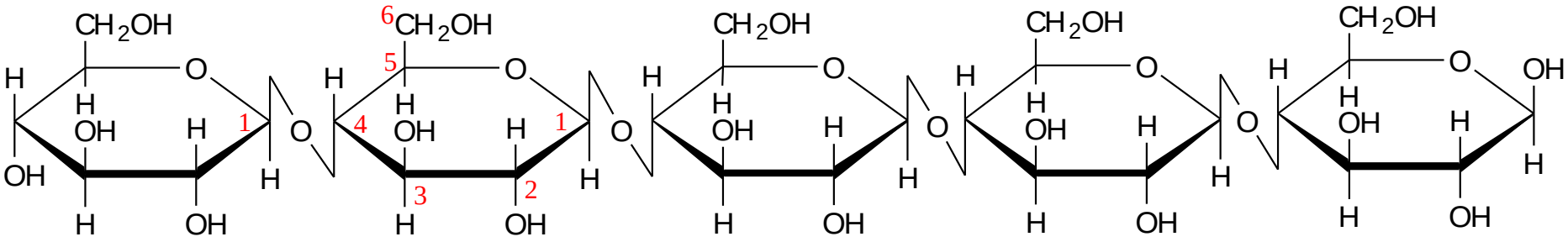


cellulose

Cellulose, a major constituent of plant cell walls, consists of long linear chains of glucose with b(1®4) linkages. Every other glucose is flipped over, due to the b linkages. This promotes intra-chain and inter-chain H-bonds and van der Waals interactions, that cause cellulose chains to be straight & rigid, and pack with a crystalline arrangement in thick bundles called microfibrils.



Schematic of arrangement of cellulose chains in a microfibril.



cellulose

Multisubunit Cellulose Synthase complexes in the plasma membrane spin out from the cell surface microfibrils consisting of 36 parallel, interacting cellulose chains.

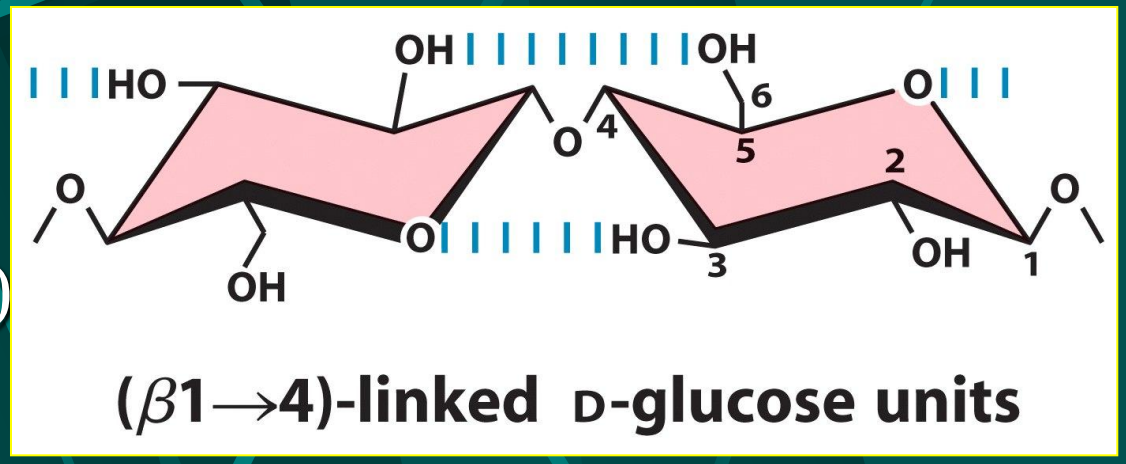
These microfibrils are very strong.

The role of cellulose is to impart strength and rigidity to plant cell walls, which can withstand high hydrostatic pressure gradients. Osmotic swelling is prevented.

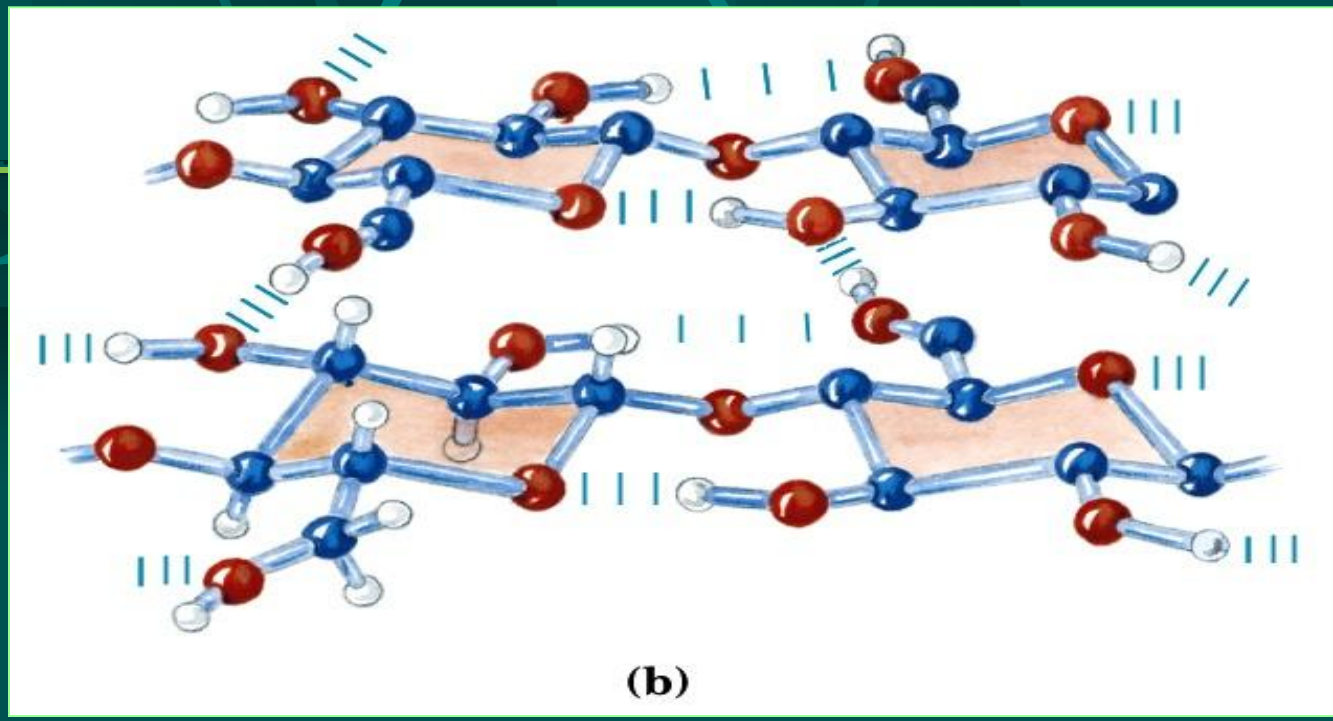
Explore and compare structures of amylose & cellulose using Chime.

Cellulose Forms Extended Sheets

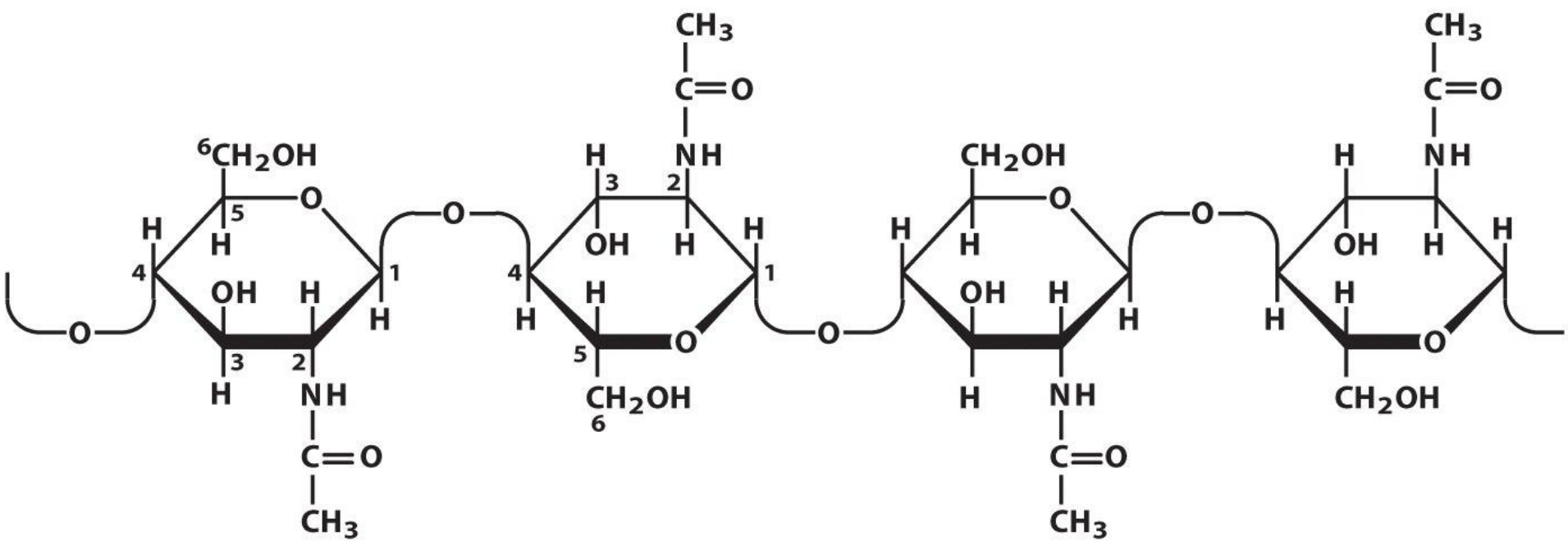
Cellulose is also pure Glucose! (but b-linkage)



Only microorganisms (and animals that host Them -- cows, termites) can digest cellulose

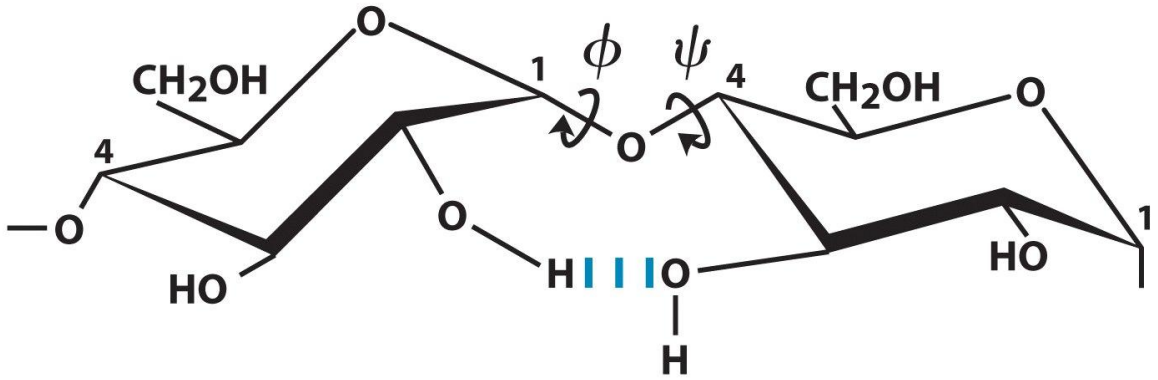


Chitin: Homopolysaccharide of GlcNac



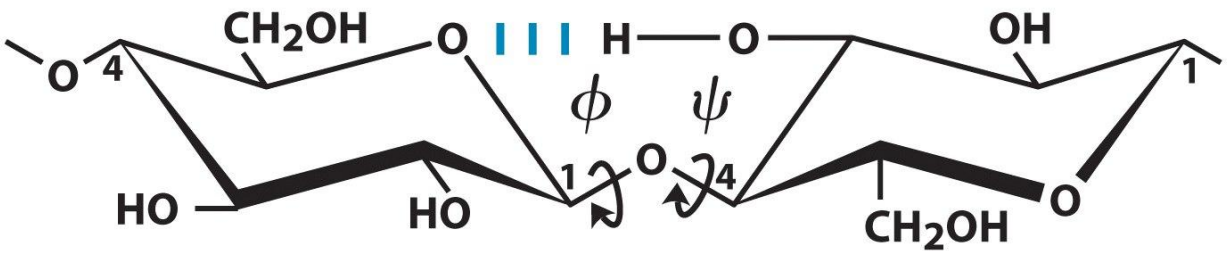
Hard sheets like cellulose: makes up exoskeletons

α vs β Linkage Strongly Influences Conformation



Amylose

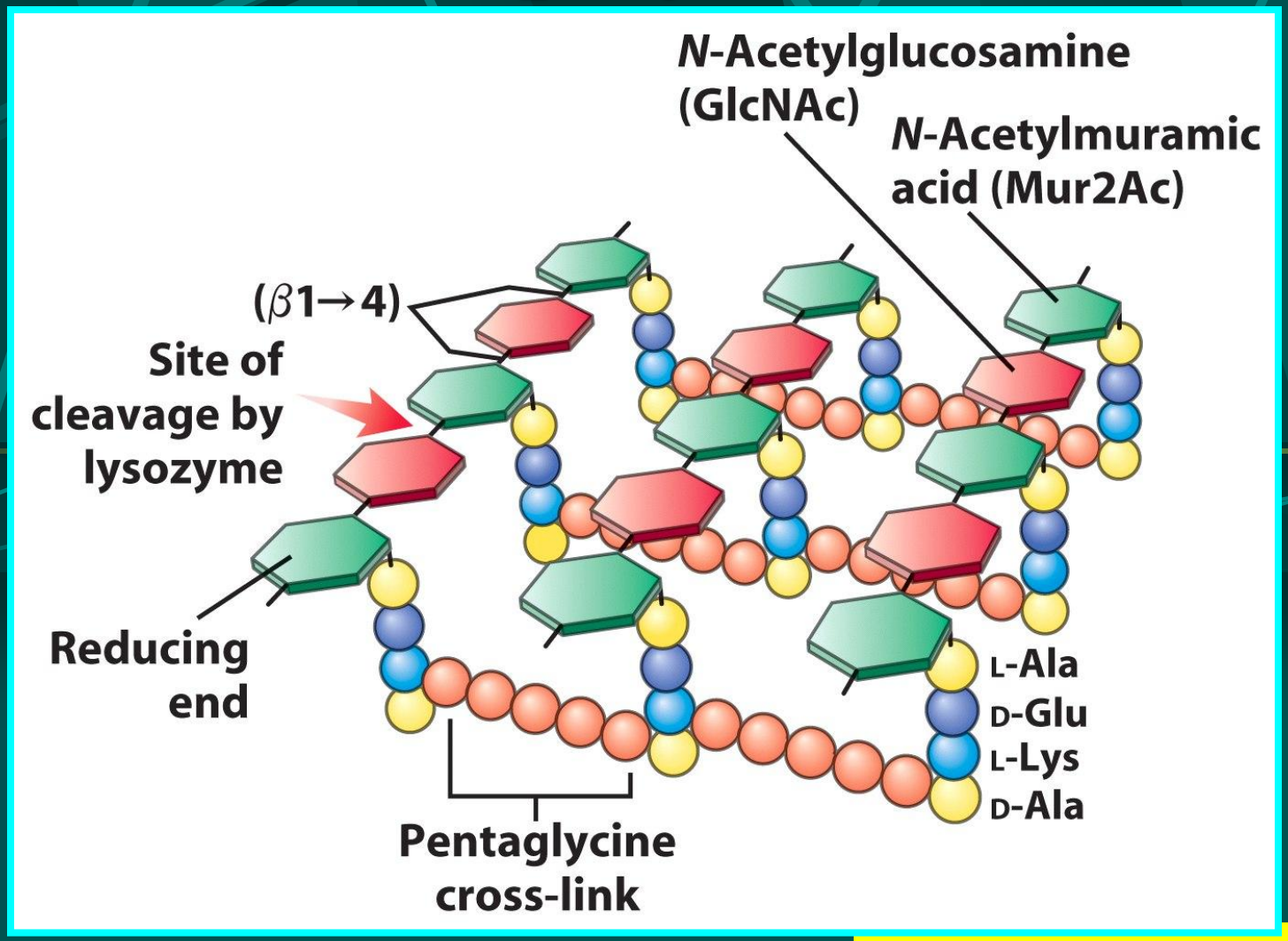
(α1→4)Glc repeats



Cellulose

(β1→4)Glc repeats

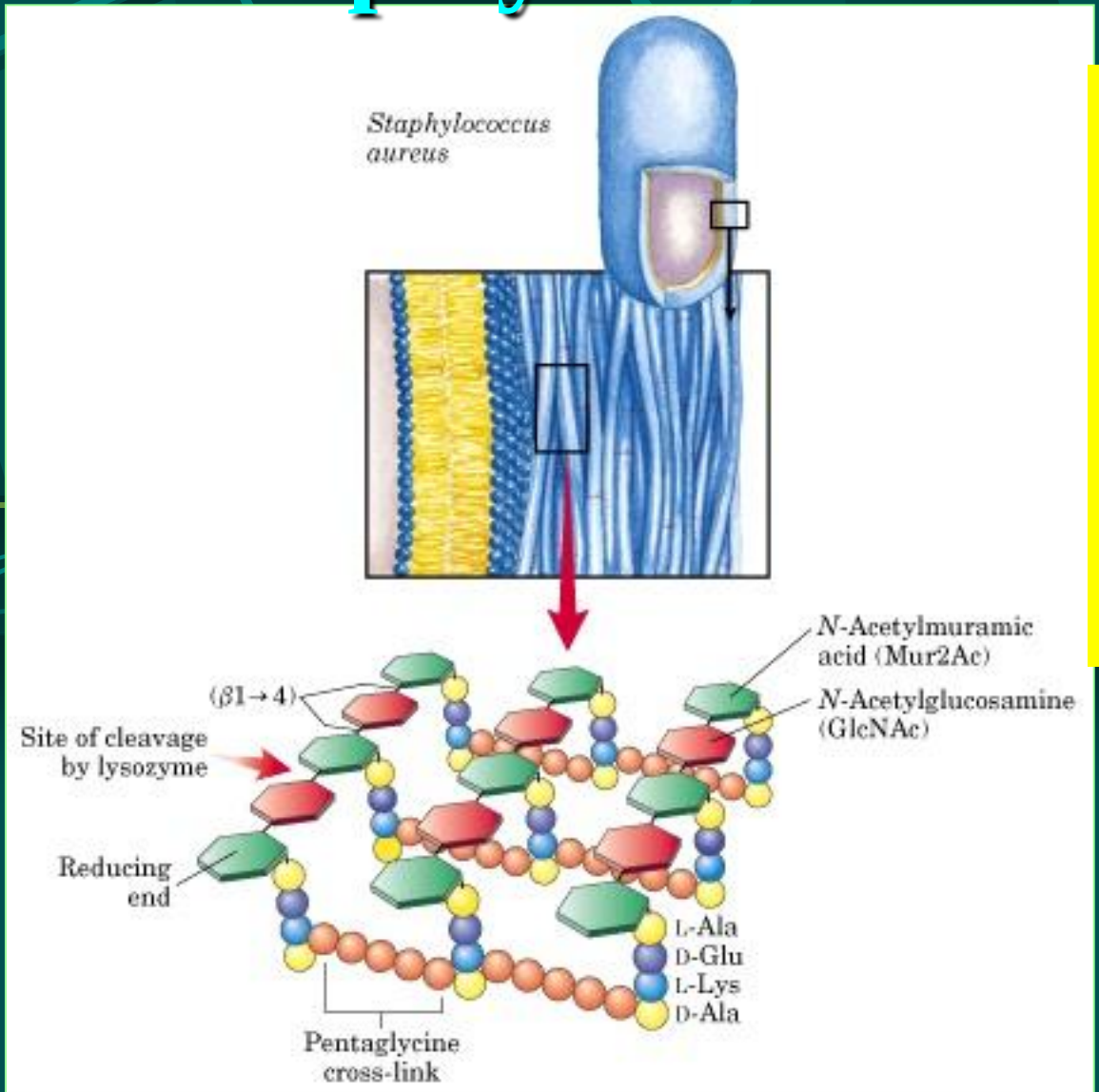
Bacterial Peptidoglycan: Simple Heteropolysaccharide



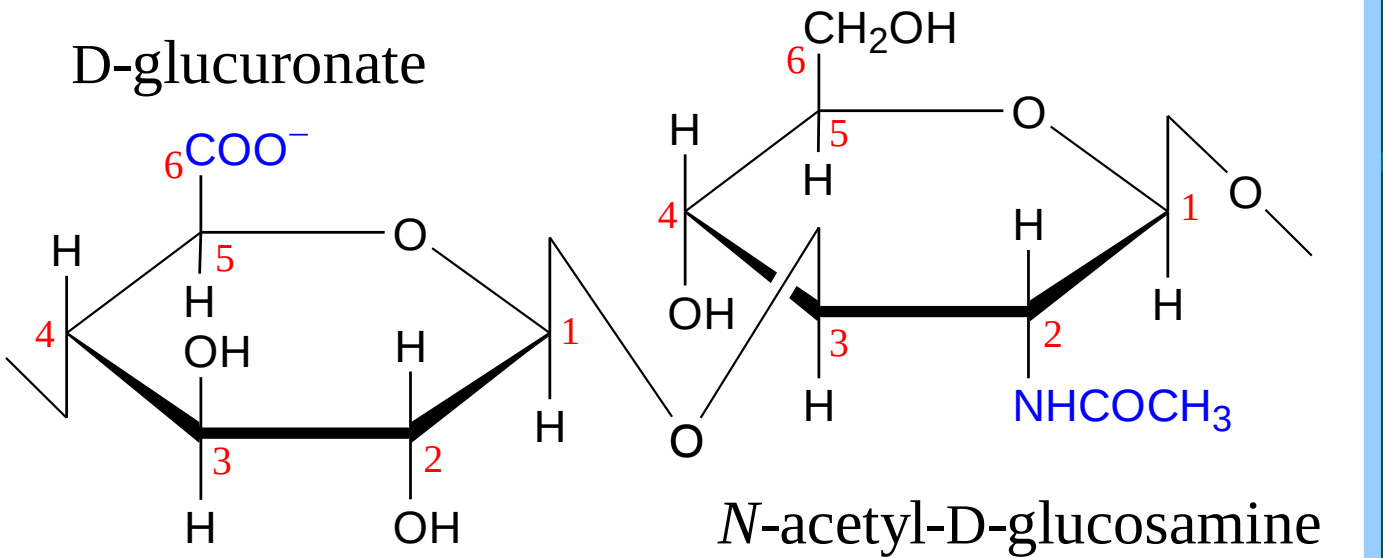
Repeating disaccharide unit

Bacterial Peptidoglycan: Simple Heteropolysaccharide

Lysozyme destroys peptidoglycan by hyrolyzing the glycosidic bond



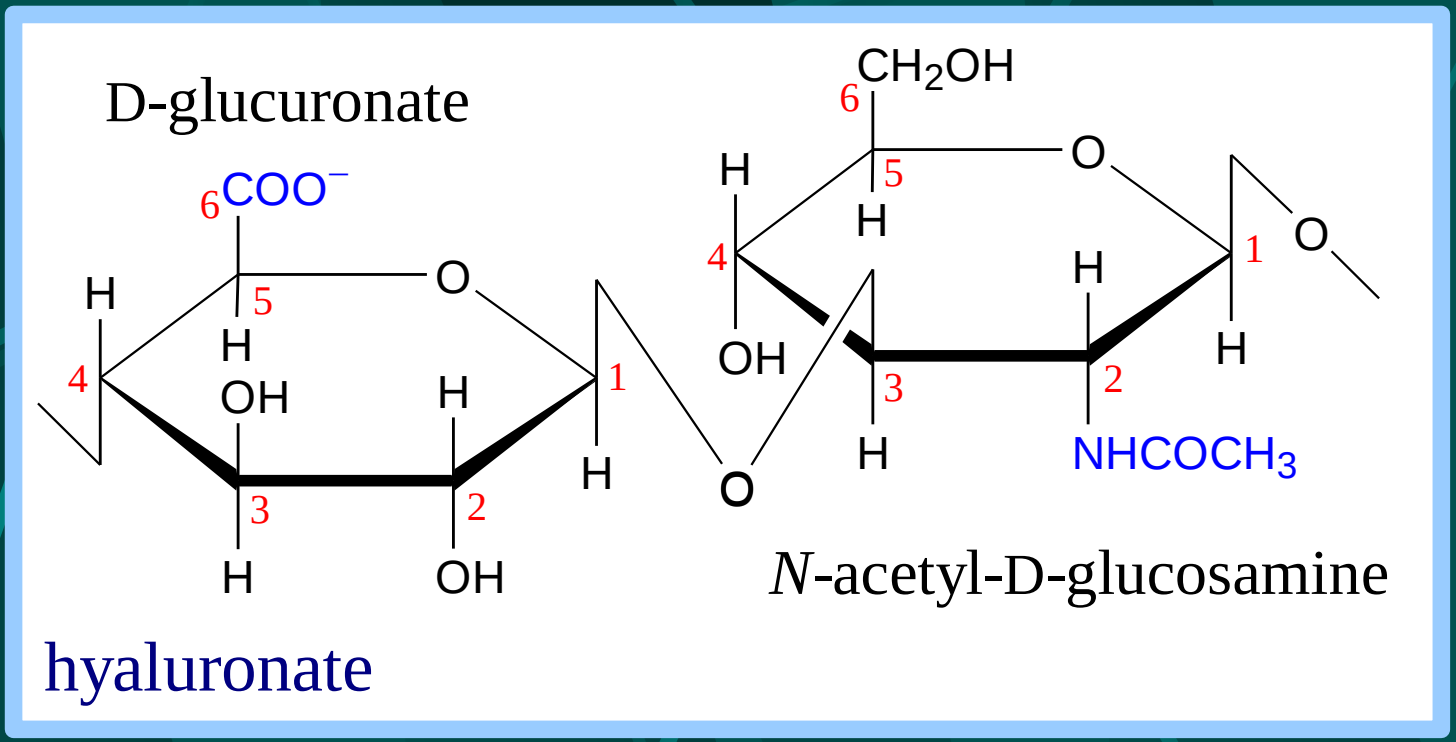
Penicillin kills bacteria by blocking synthesis of peptide crosslinks



Glycosaminoglycans (mucopolysaccharides) are polymers of **repeating disaccharides**.

Within the disaccharides, the sugars tend to be modified, with acidic groups, amino groups, sulfated hydroxyl and amino groups, etc.

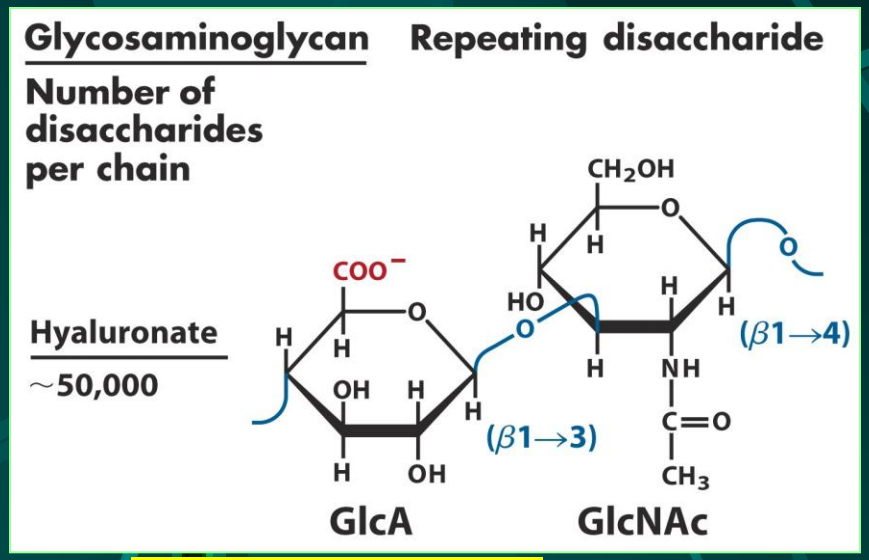
Glycosaminoglycans tend to be **negatively charged**, because of the prevalence of acidic groups.



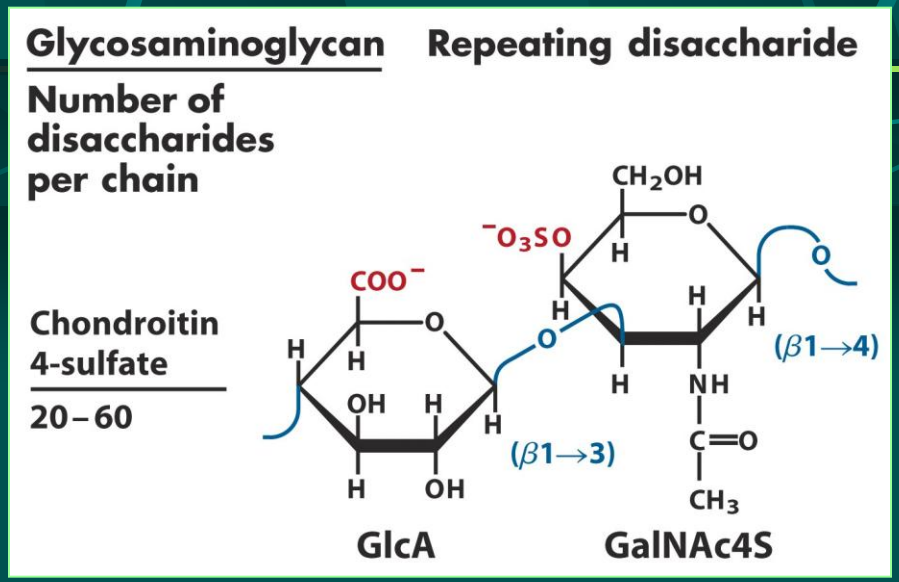
Hyaluronate is a glycosaminoglycan with a repeating disaccharide consisting of 2 glucose derivatives, glucuronate (glucuronic acid) & *N*-acetyl-glucosamine.

The glycosidic linkages are b(1®3) & b(1®4).

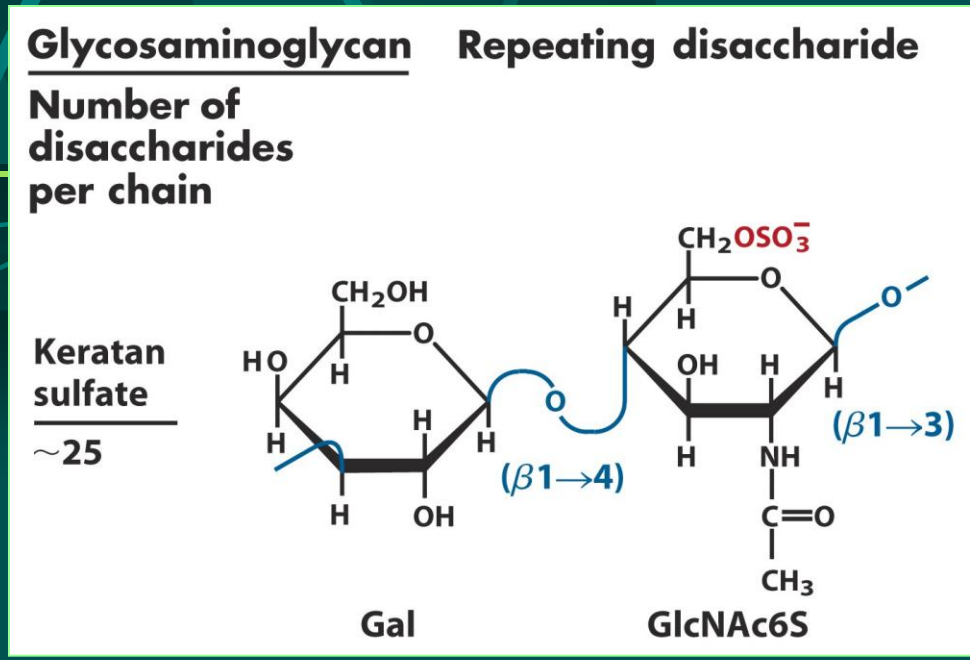
Glycosamino-glycans of the Extracellular Matrix



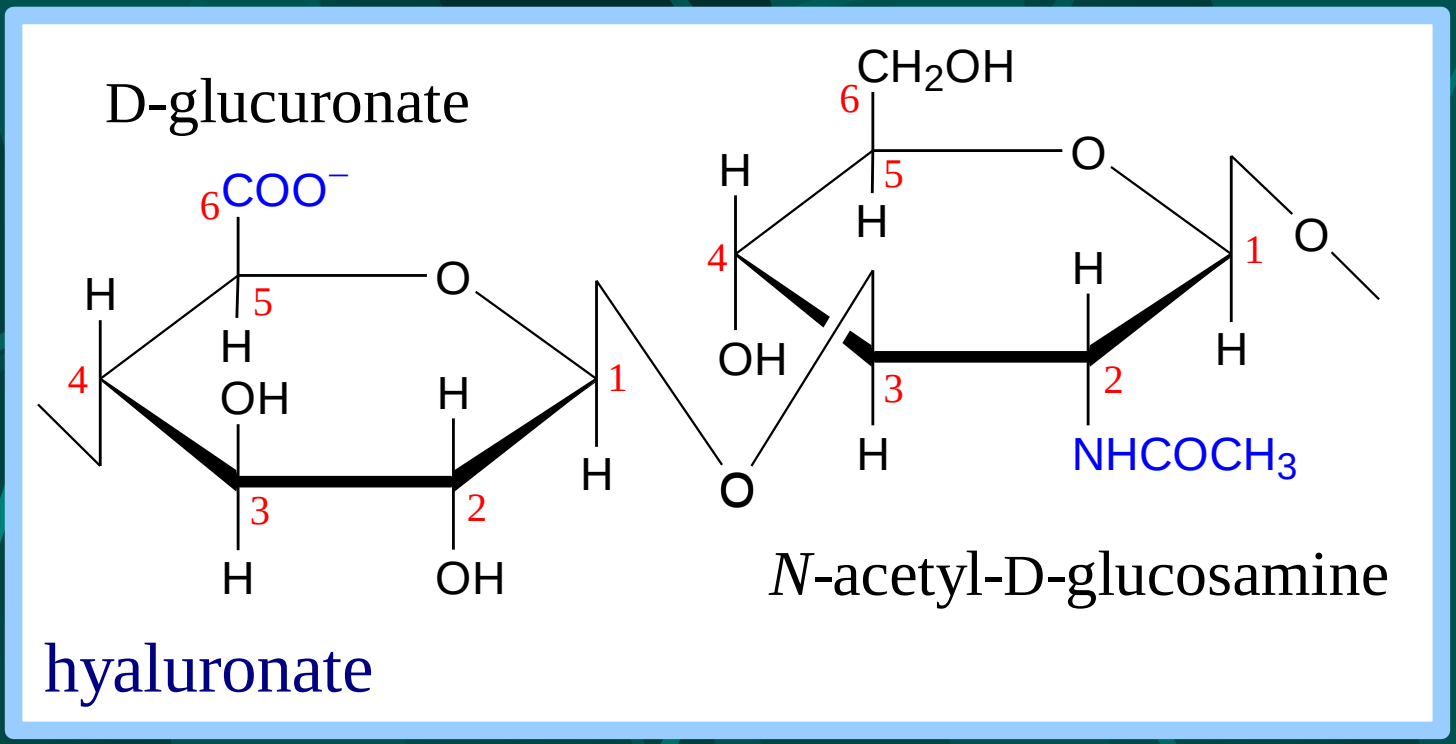
Joints and eye



Cartilage, tendons, ligaments



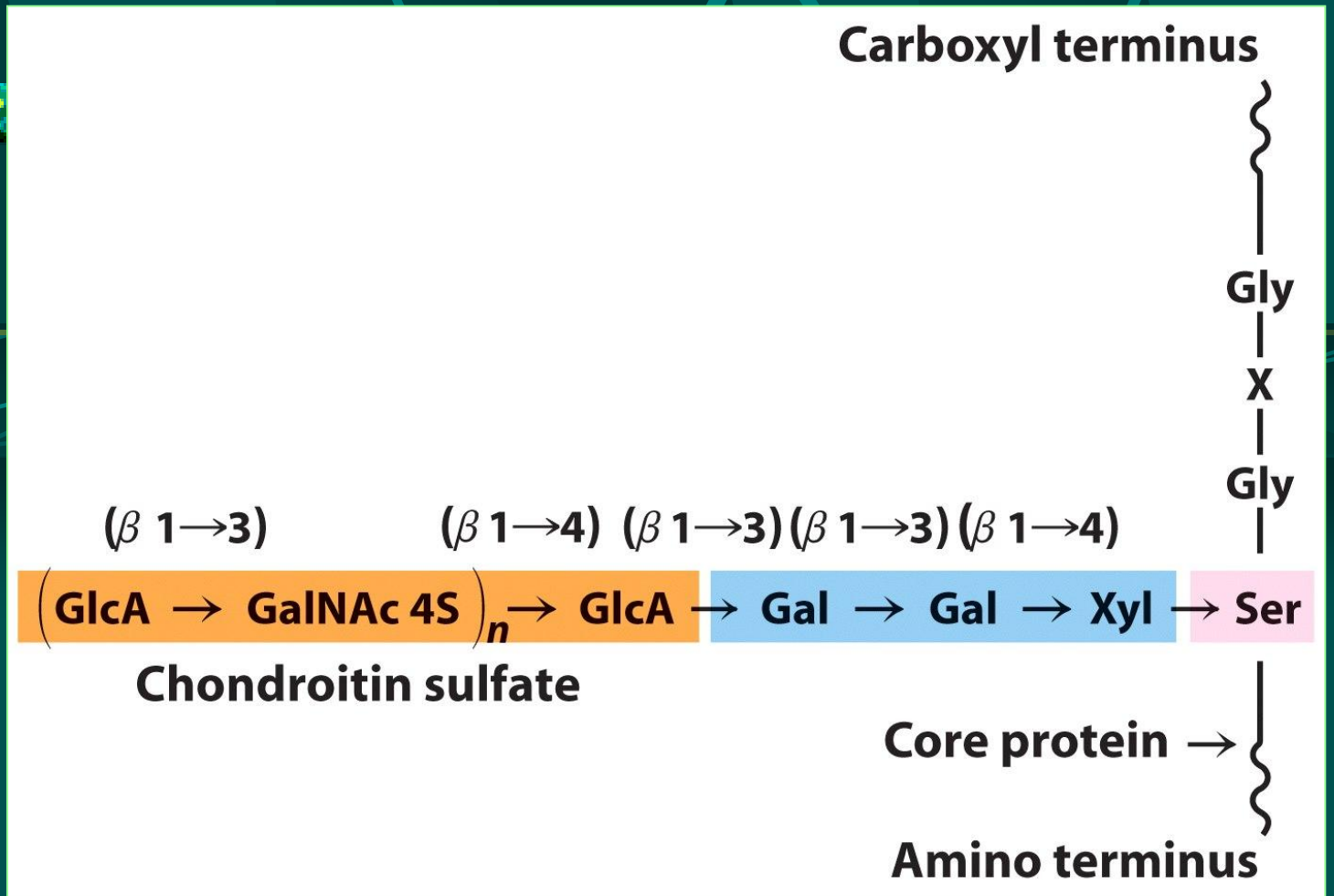
Cartilage, bone, horny structures

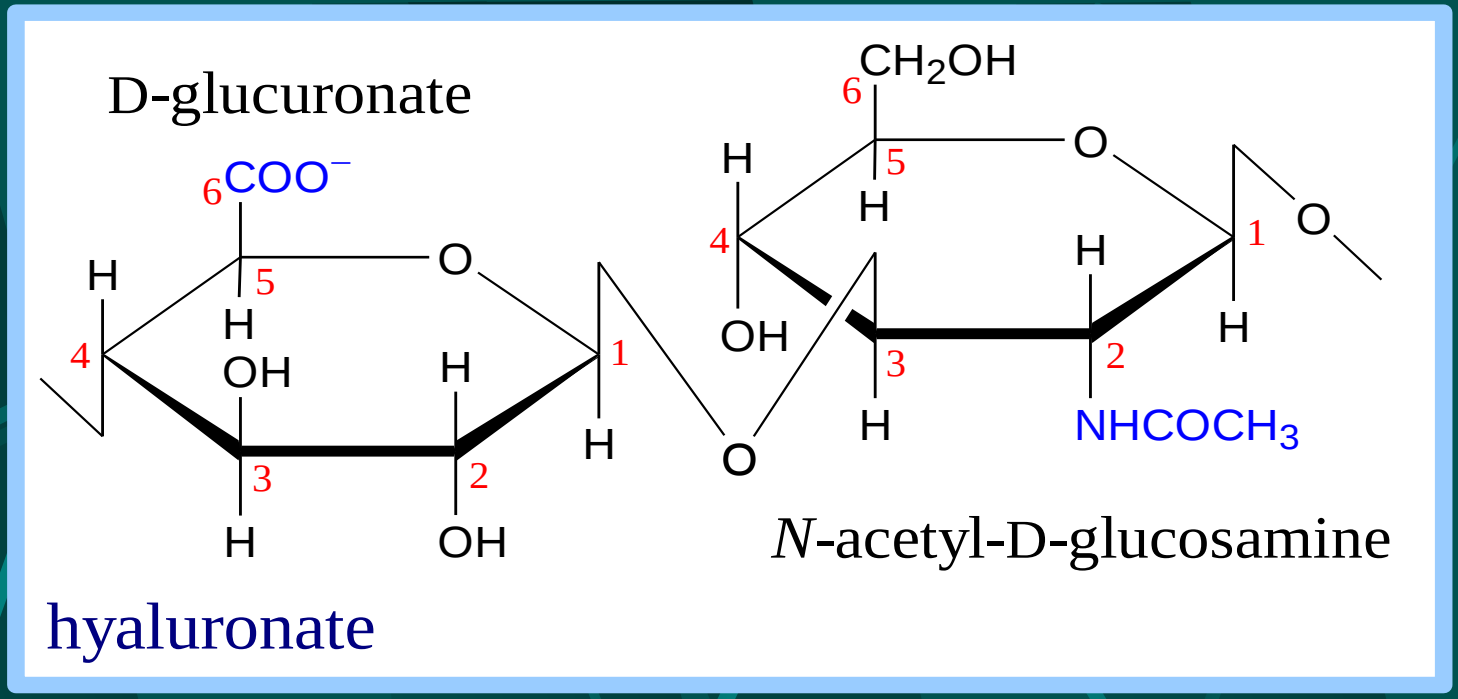


Proteoglycans are glycosaminoglycans that are covalently linked to specific **core proteins**.

Some proteoglycans of the extracellular matrix in turn link non-covalently to hyaluronate via protein domains called **link modules**.

Proteoglycan: Glycosaminoglycan Linked to Proteins



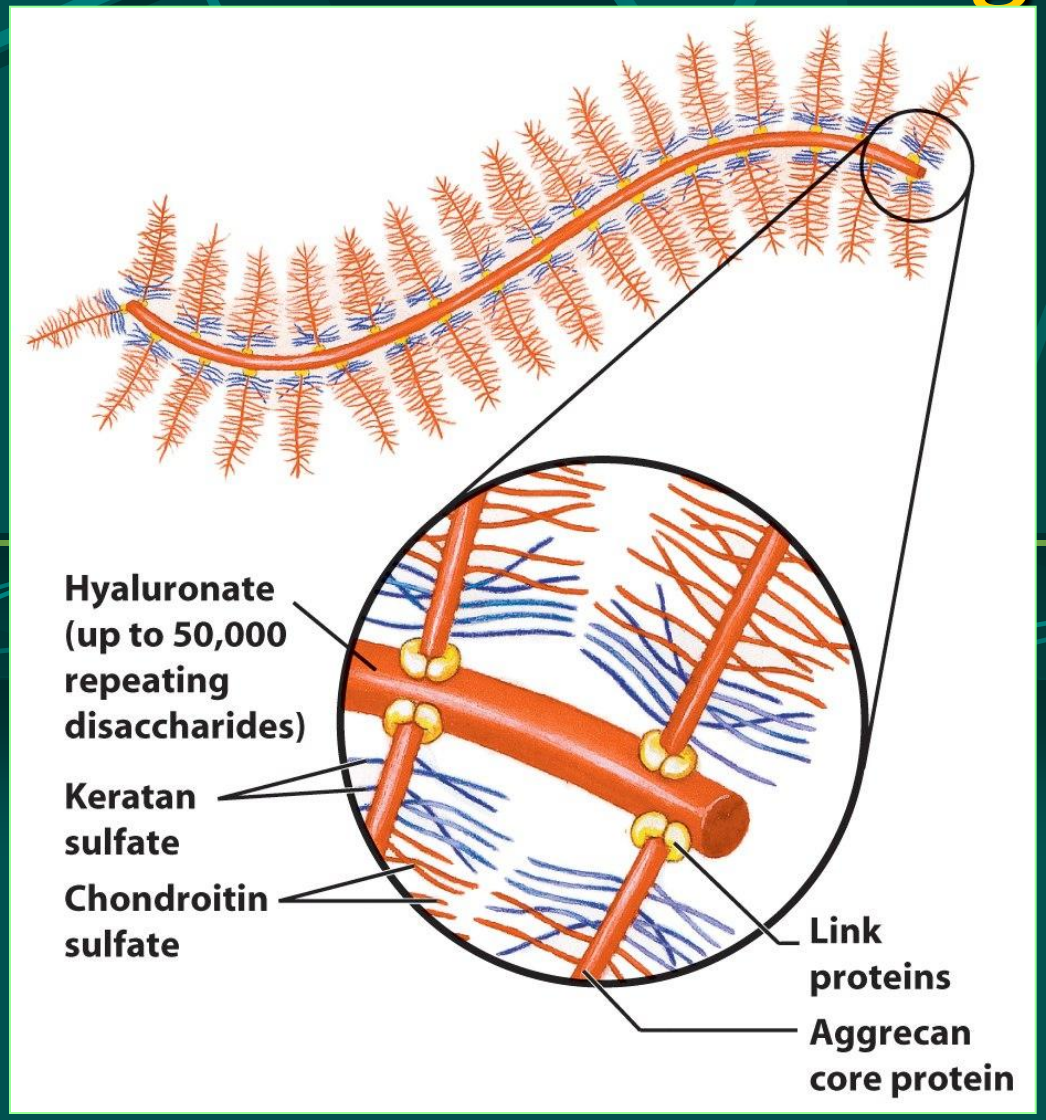


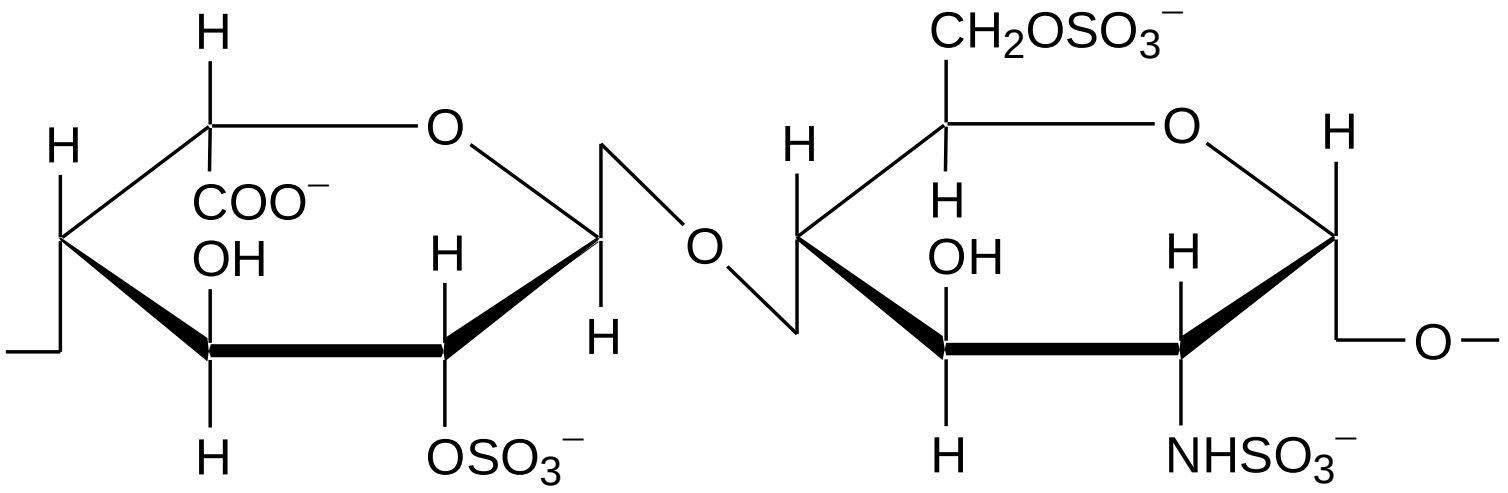
For example, in cartilage multiple copies of the **aggrecan** proteoglycan bind to an extended hyaluronate backbone to form a large complex.

Versican, another proteoglycan that binds to hyaluronate, is in the extracellular matrix of loose connective tissues.

See web sites on aggrecan and aggrecan plus versican.

Proteoglycan Aggregate: Resilient Structure of Cartilage





heparin or heparan sulfate - examples of residues

Heparan sulfate is initially synthesized on a membrane-embedded core protein as a polymer of alternating ***N*-acetylglucosamine** and **glucuronate** residues.

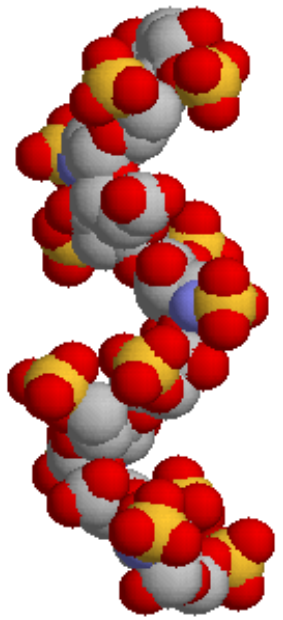
Later, in segments of the polymer, glucuronate residues may be converted to the sulfated sugar **iduronic acid**, while *N*-acetylglucosamine residues may be deacetylated and/or sulfated.

Heparin, a soluble glycosaminoglycan found in granules of mast cells, has a structure similar to that of heparan sulfates, but is more **highly sulfated**.

When released into the blood, it inhibits clot formation by interacting with the protein antithrombin.

Heparin has an **extended helical conformation**.

PDB 1RID

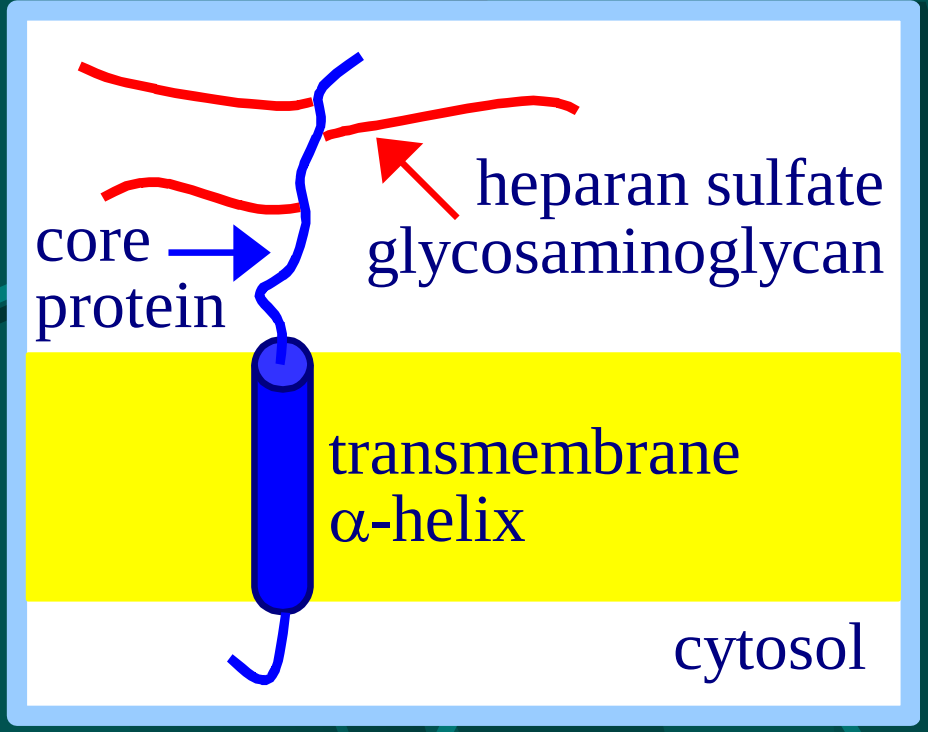


heparin: $(IDS\text{-}SGN)_5$

C O N S

Charge repulsion by the many negatively charged groups may contribute to this conformation.

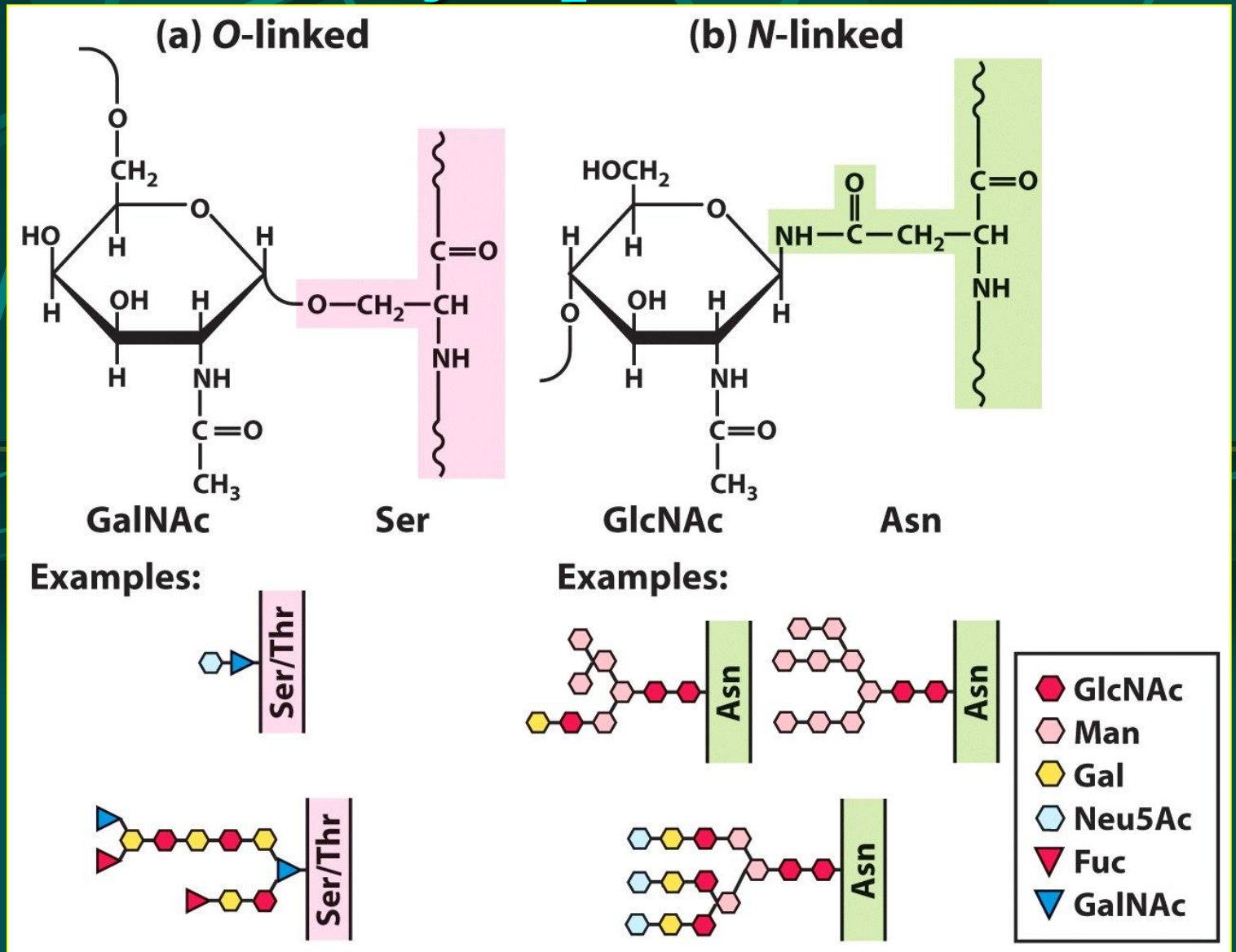
Heparin shown has 10 residues, alternating IDS (iduronate-2-sulfate) & SGN (N-sulfo-glucosamine-6-sulfate).



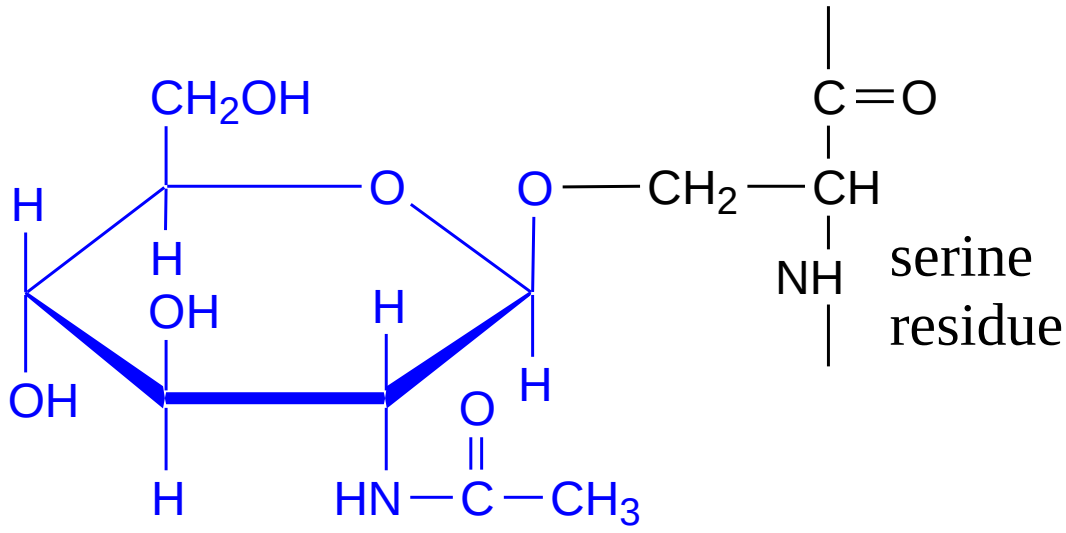
Some cell surface **heparan sulfate** glycosaminoglycans remain covalently linked to core proteins embedded in the plasma membrane.

Proteins involved in signaling & adhesion at the cell surface recognize and bind segments of heparan sulfate chains having particular patterns of sulfation.

Oligosaccharide Linkages in Glycoproteins



Oligosaccharides that are covalently attached to proteins or to membrane lipids may be linear or branched chains.

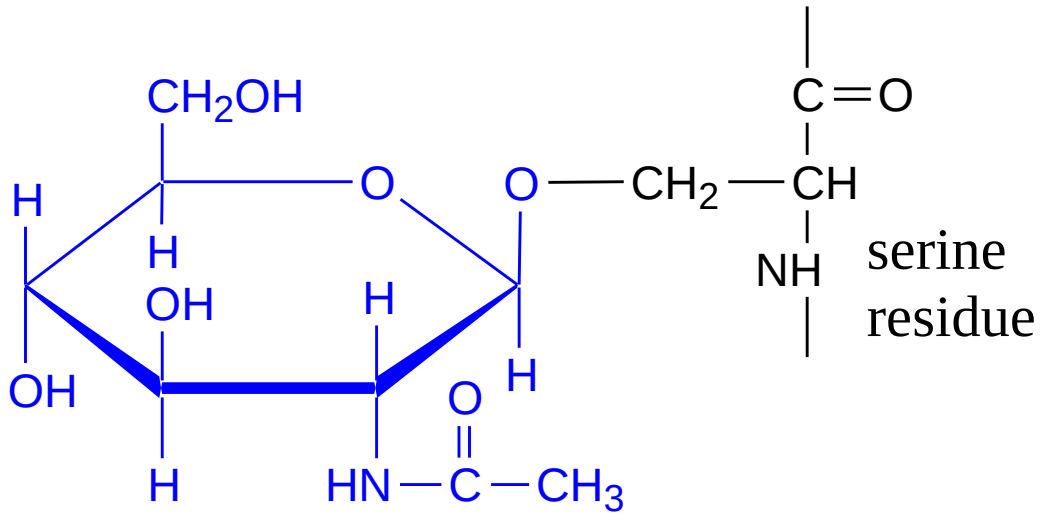


β-D-*N*-acetylglucosamine

O-linked oligosaccharide chains of glycoproteins vary in complexity.

They link to a protein via a glycosidic bond between a sugar residue & a **serine or threonine OH**.

O-linked oligosaccharides have roles in **recognition**, **interaction**, and **enzyme regulation**.

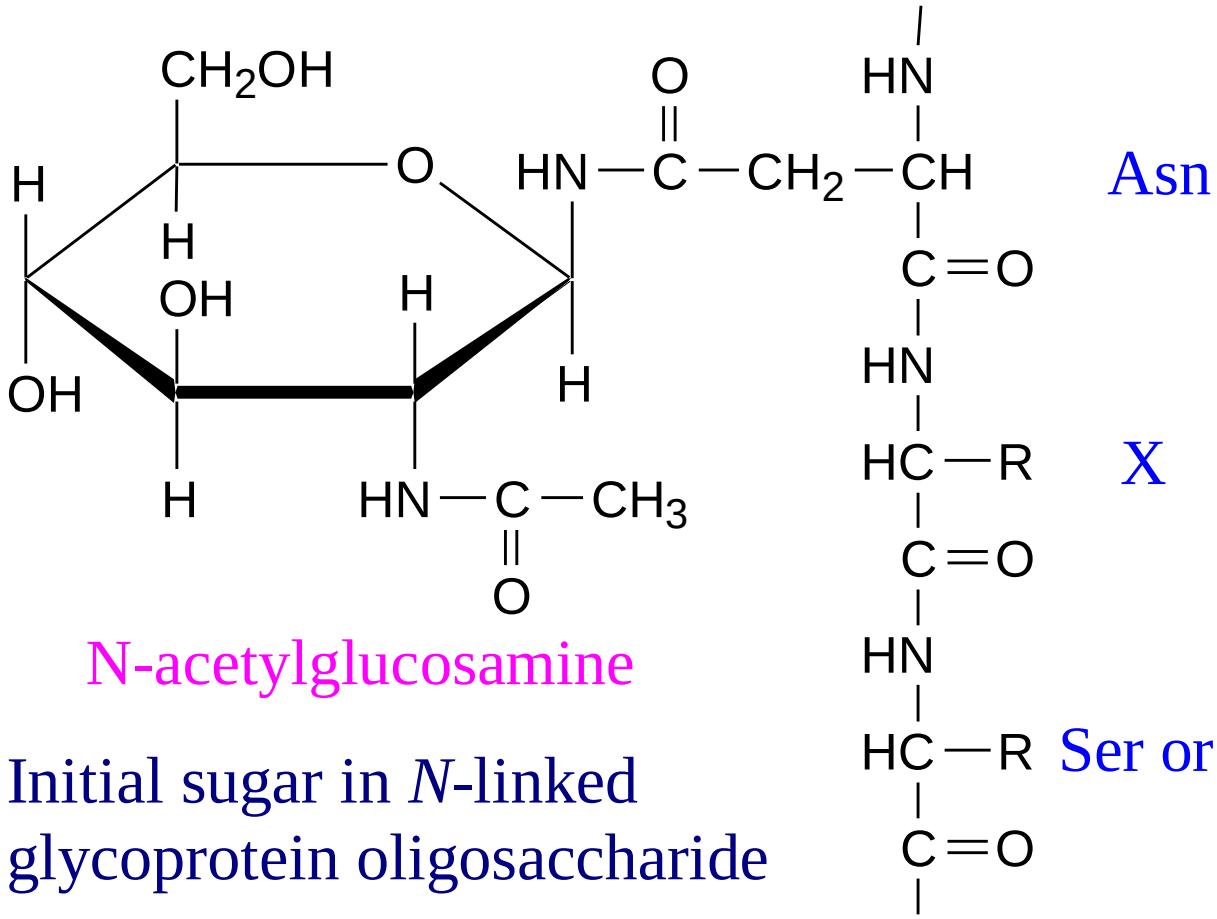


β-D-*N*-acetylglucosamine

***N*-acetylglucosamine** (**GlcNAc**) is a common O-linked glycosylation of protein serine or threonine residues.

Many cellular proteins, including enzymes & transcription factors, are **regulated** by reversible GlcNAc attachment.

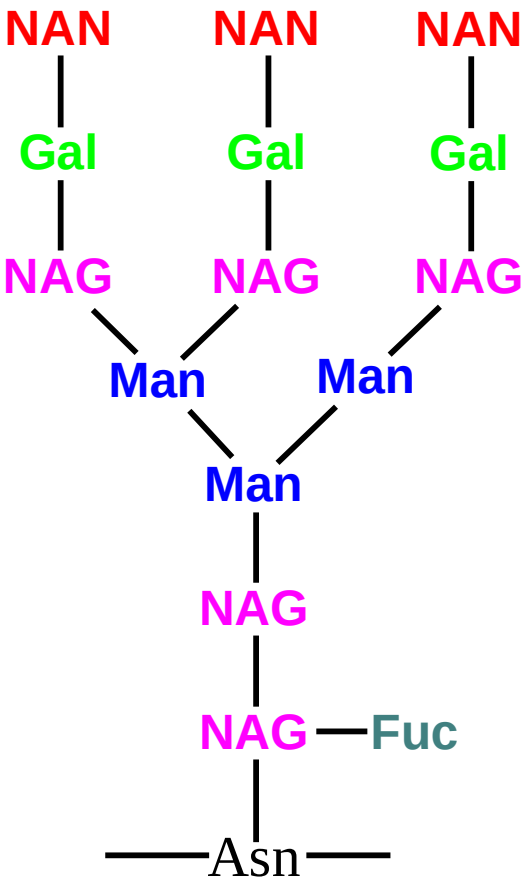
Often attachment of GlcNAc to a protein OH **alternates with phosphorylation**, with these 2 modifications having opposite regulatory effects (stimulation or inhibition).



N-acetylglucosamine

Initial sugar in *N*-linked glycoprotein oligosaccharide

***N*-linked oligosaccharides** of glycoproteins tend to be complex and branched. First ***N*-acetylglucosamine** is linked to a protein via the side-chain N of an asparagine residue in a particular 3-amino acid sequence.



***N*-linked oligosaccharide**

Key:
NAN = N-acetylneuraminate
Gal = galactose
NAG = N-acetylglucosamine
Man = mannose
Fuc = fucose

Additional monosaccharides are added, and the *N*-linked oligosaccharide chain is modified by removal and addition of residues, to yield a characteristic branched structure.

Many proteins secreted by cells have attached N-linked oligosaccharide chains.

Genetic diseases have been attributed to deficiency of particular enzymes involved in synthesizing or modifying oligosaccharide chains of these glycoproteins.

Such diseases, and gene knockout studies in mice, have been used to define pathways of modification of oligosaccharide chains of glycoproteins and glycolipids.

Carbohydrate chains of plasma membrane glycoproteins and glycolipids usually face the outside of the cell.

They have roles in cell-cell interaction and signaling, and in forming a protective layer on the surface of some cells.

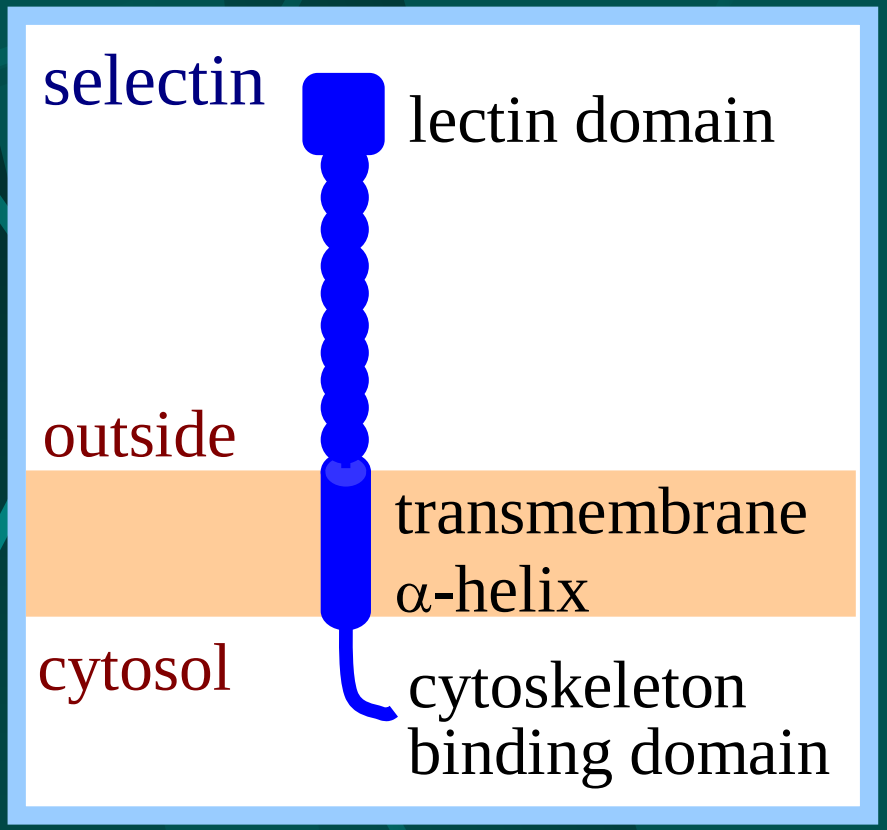
Lectins are glycoproteins that recognize and bind to specific oligosaccharides. A few examples:

- **Concanavalin A** and **wheat germ agglutinin** are plant lectins that have been useful research tools.
- **Mannan-binding lectin** (**MBL**) is a glycoprotein found in blood plasma.

  It associates with cell surface carbohydrates of disease-causing microorganisms, promoting phagocytosis of these organisms as part of the immune response.

Selectins are integral proteins of mammalian cell plasma membranes with roles in **cell-cell recognition** & binding.

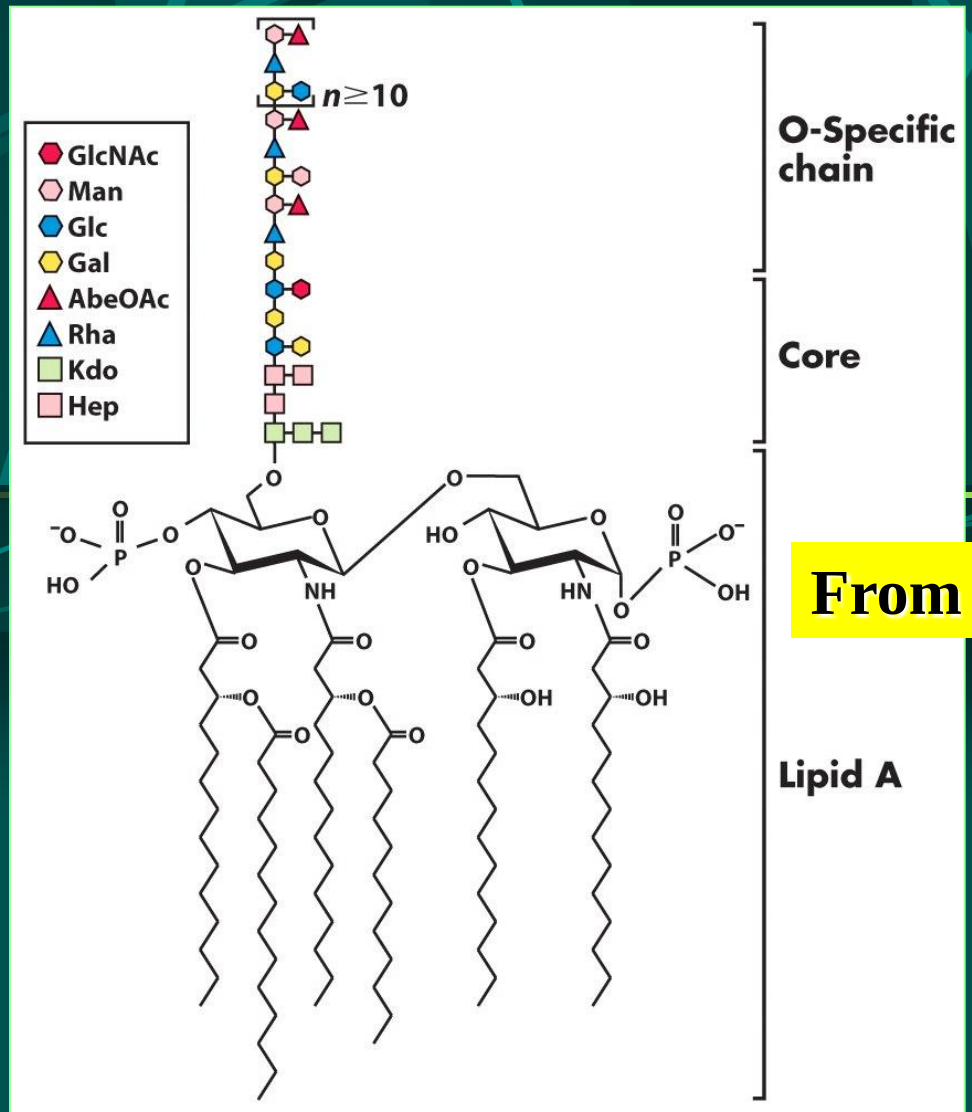
A **lectin-like domain** is at the end of an extracellular segment that extends out from the cell surface.

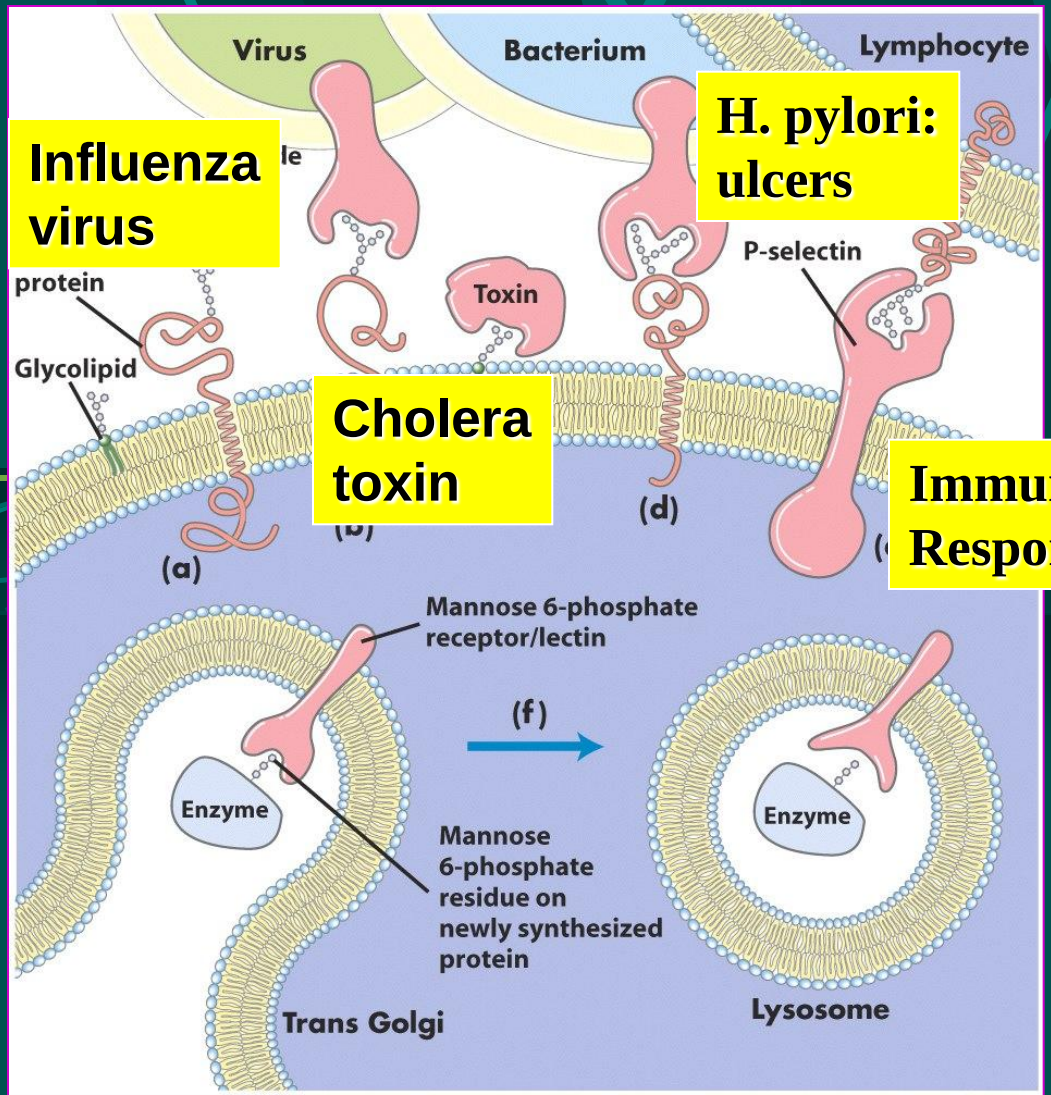


A **cleavage site** just outside the transmembrane a-helix provides a mechanism for regulated release of some lectins from the cell surface.

A **cytosolic domain** participates in regulated interaction with the actin cytoskeleton.

Complex Polysaccharides Are Also Added to Lipids



From bacterial outer membrane

Oligosaccharides in Recognition Processes