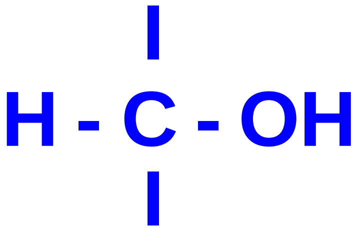
Carbohydrates

Carbohydrates (glycans) are polyhydroxy aldehyde or ketones with following basic composition:

$(CH_2O)_n$ or

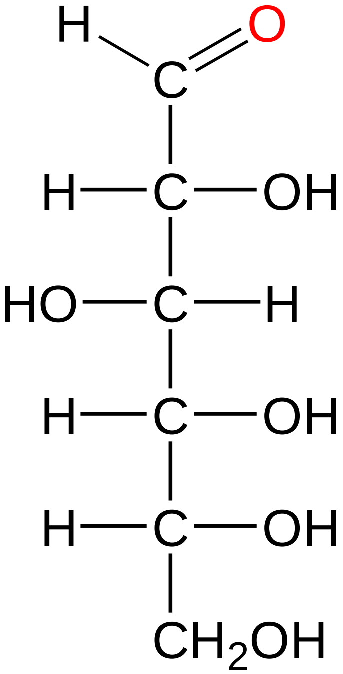


$$H-\overset{|}{\underset{|}{C}}-OH$$

- 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.

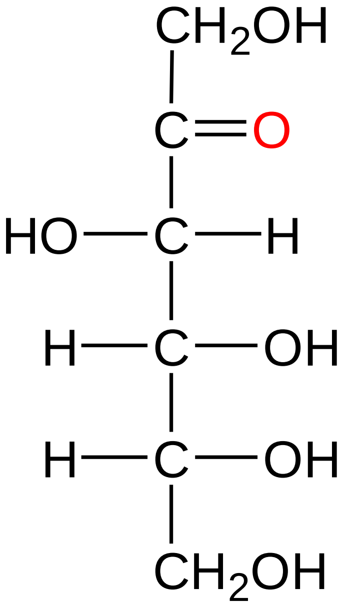
Monosaccharides

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



D-glucose

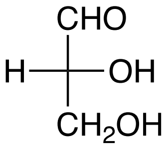
Ketoses (e.g., fructose) have a keto group, usually at C2.



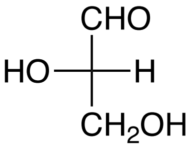
D-fructose

The Aldotetroses. Glyceraldehyde is the simplest carbohydrate (C_3, aldotriose, 2,3-dihydroxypropanal). The next carbohydrate are aldotetroses (C_4, 2,3,4-trihydroxybutanal).

aldotriose



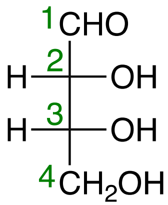
D-glyceraldehyde



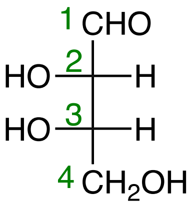
L-glyceraldehyde

aldotetroses

highest numbered "chiral" carbon



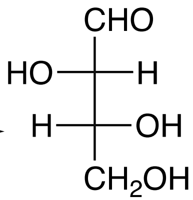
D-erythrose



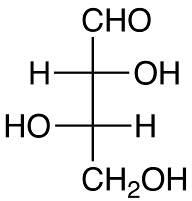
L-erythrose

highest numbered "chiral" carbon

highest numbered "chiral" carbon



D-threose

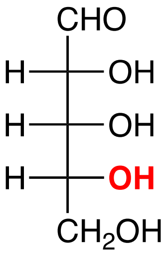


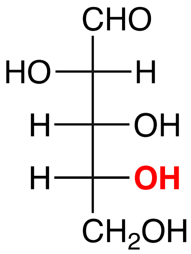
L-threose

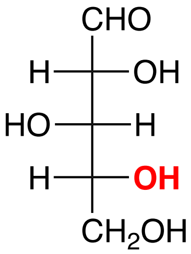
highest numbered "chiral" carbon

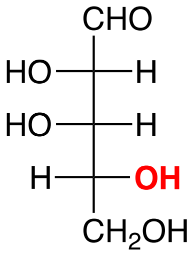
Aldopentoses and Aldohexoses.

Aldopentoses: C_5, three chiral carbons, eight stereoisomers

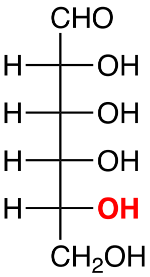

D-ribose

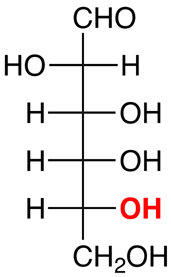

D-arabinose

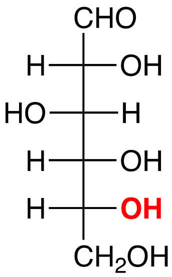

D-xylose

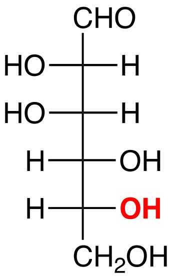

D-lyxose

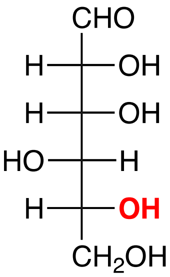
Aldohexoses: C_6, four chiral carbons, sixteen stereoisomers

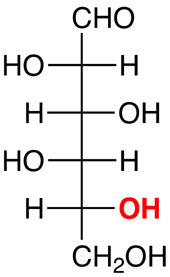

D-allose

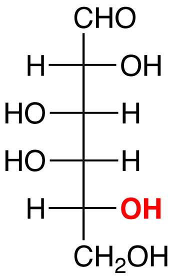

D-altrose

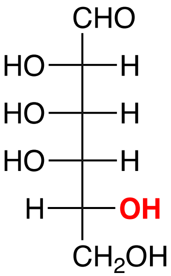

D- glucose


D-mannose


D-gulose


D-idose

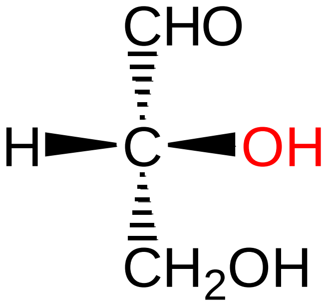

D-galactose


D-talose

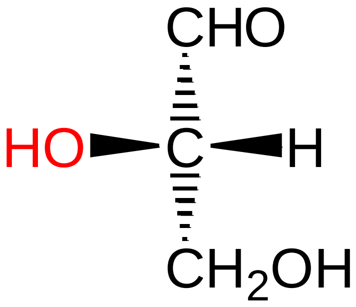
D vs L Designation

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

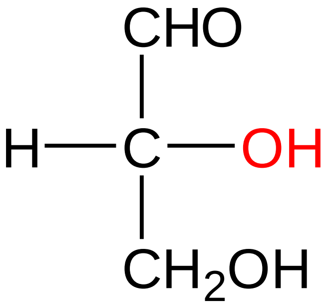
The lower representations are Fischer Projections.



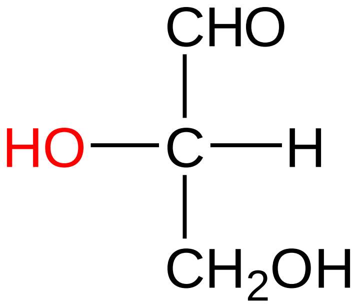
D-glyceraldehyde



L-glyceraldehyde



D-glyceraldehyde

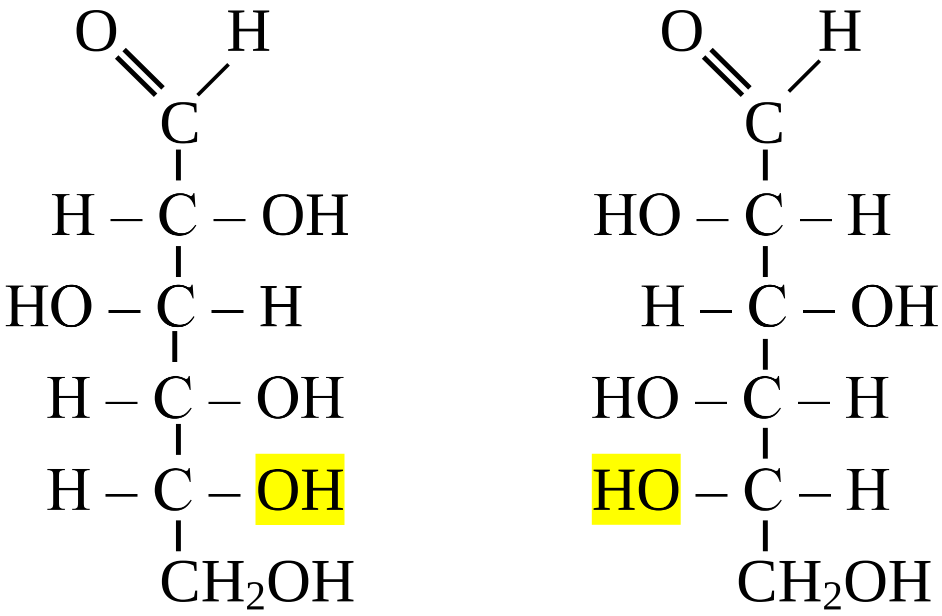


L-glyceraldehyde

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.



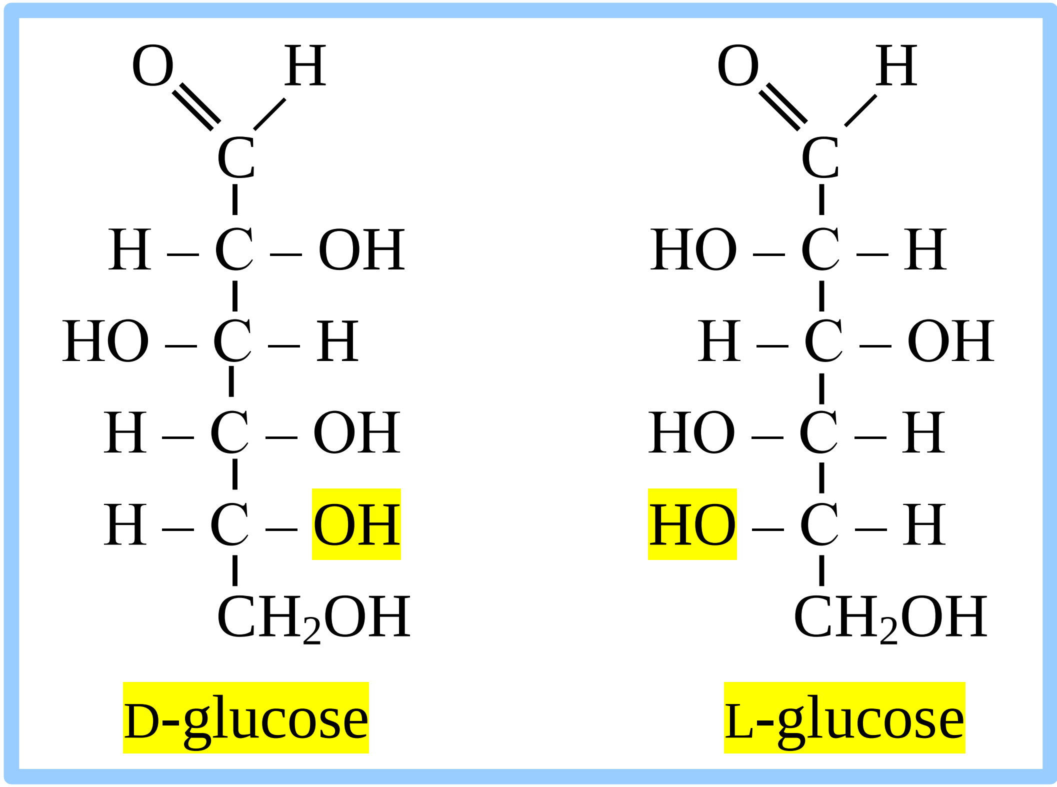
D-glucose

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

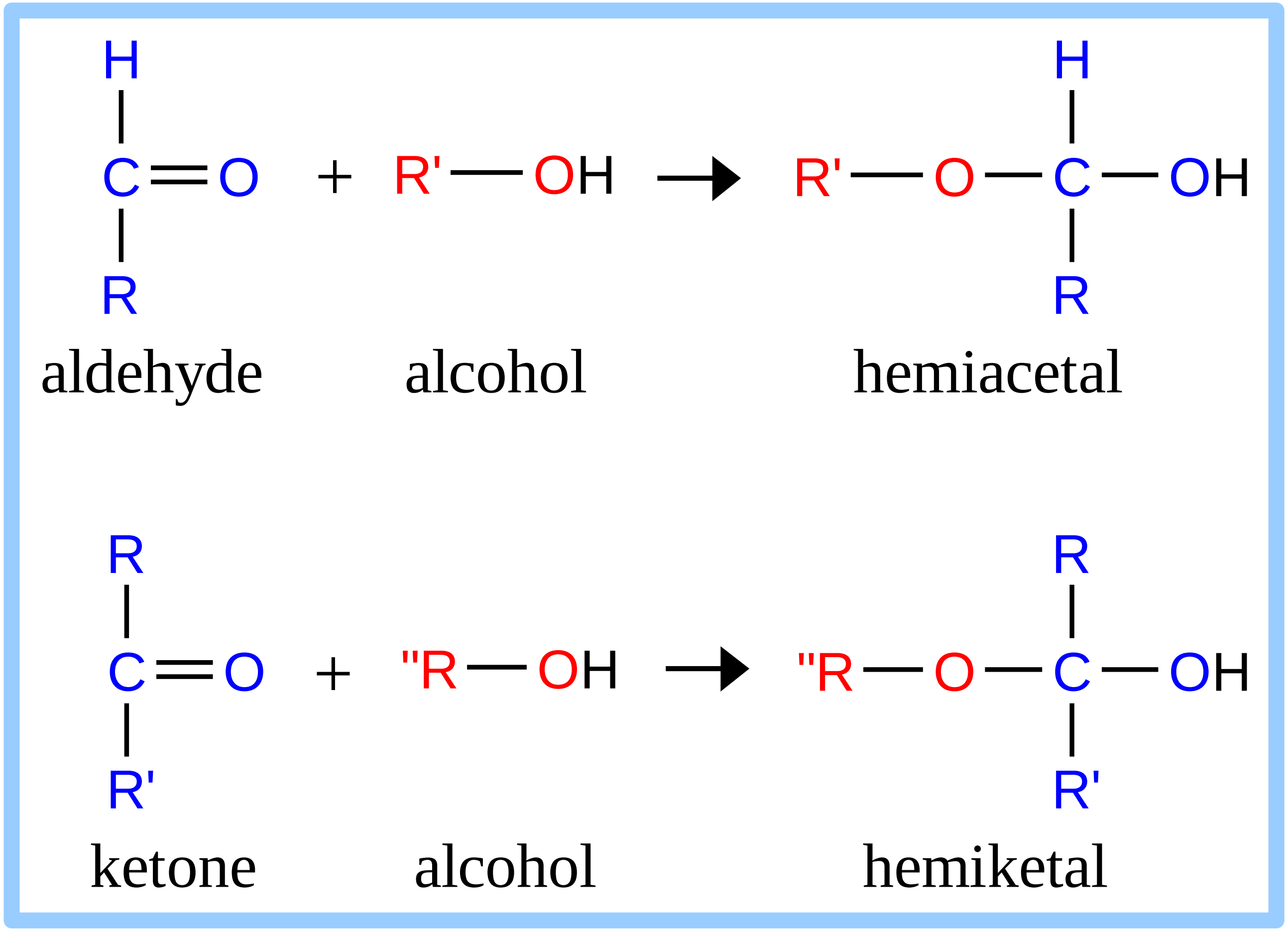
The number of stereoisomers is $\mathbf{2^n}$, where n is the number of asymmetric centers.

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

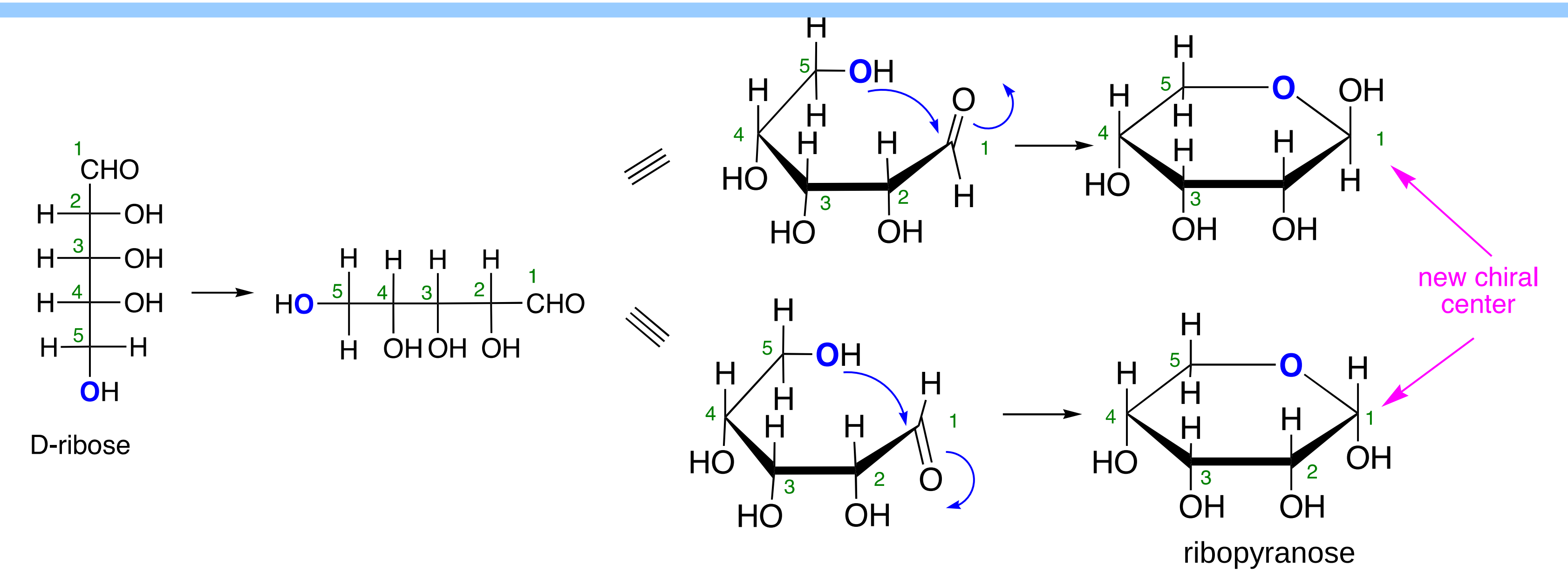
Hemiacetal & hemiketal formation

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

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



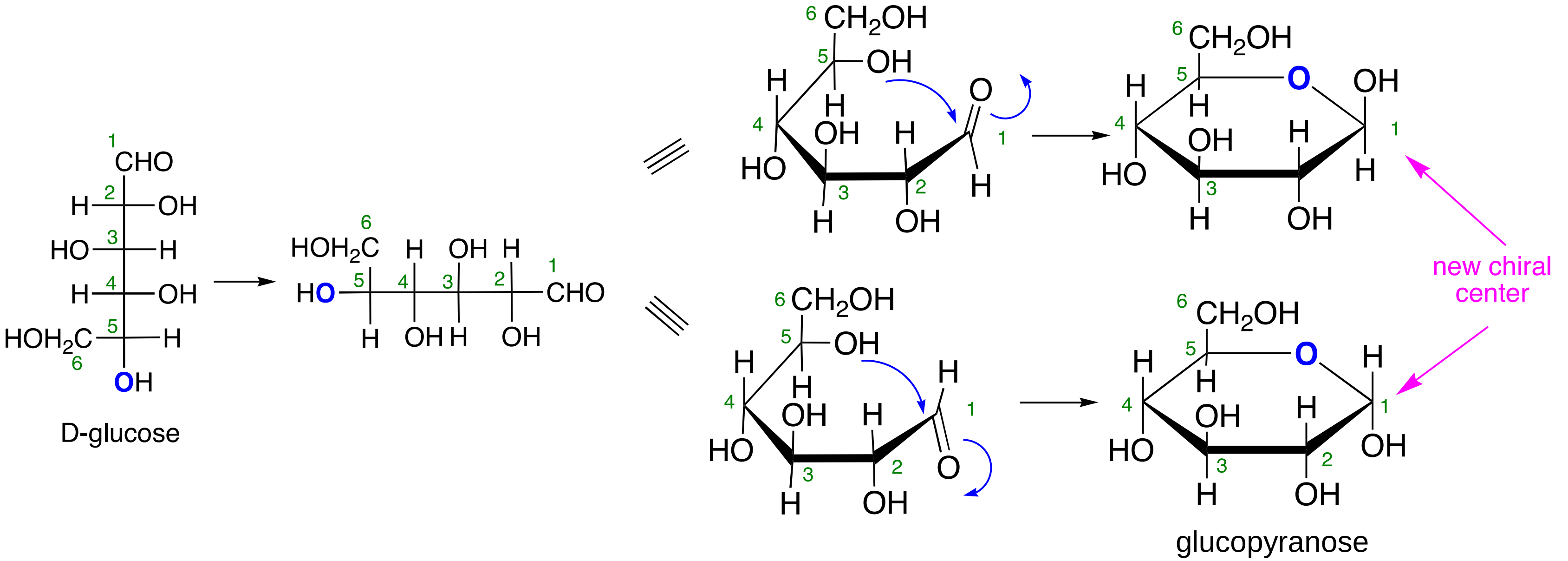
Cyclic Forms of Carbohydrates: Pyranose Forms.



D-ribose

ribopyranose

new chiral center



D-glucose

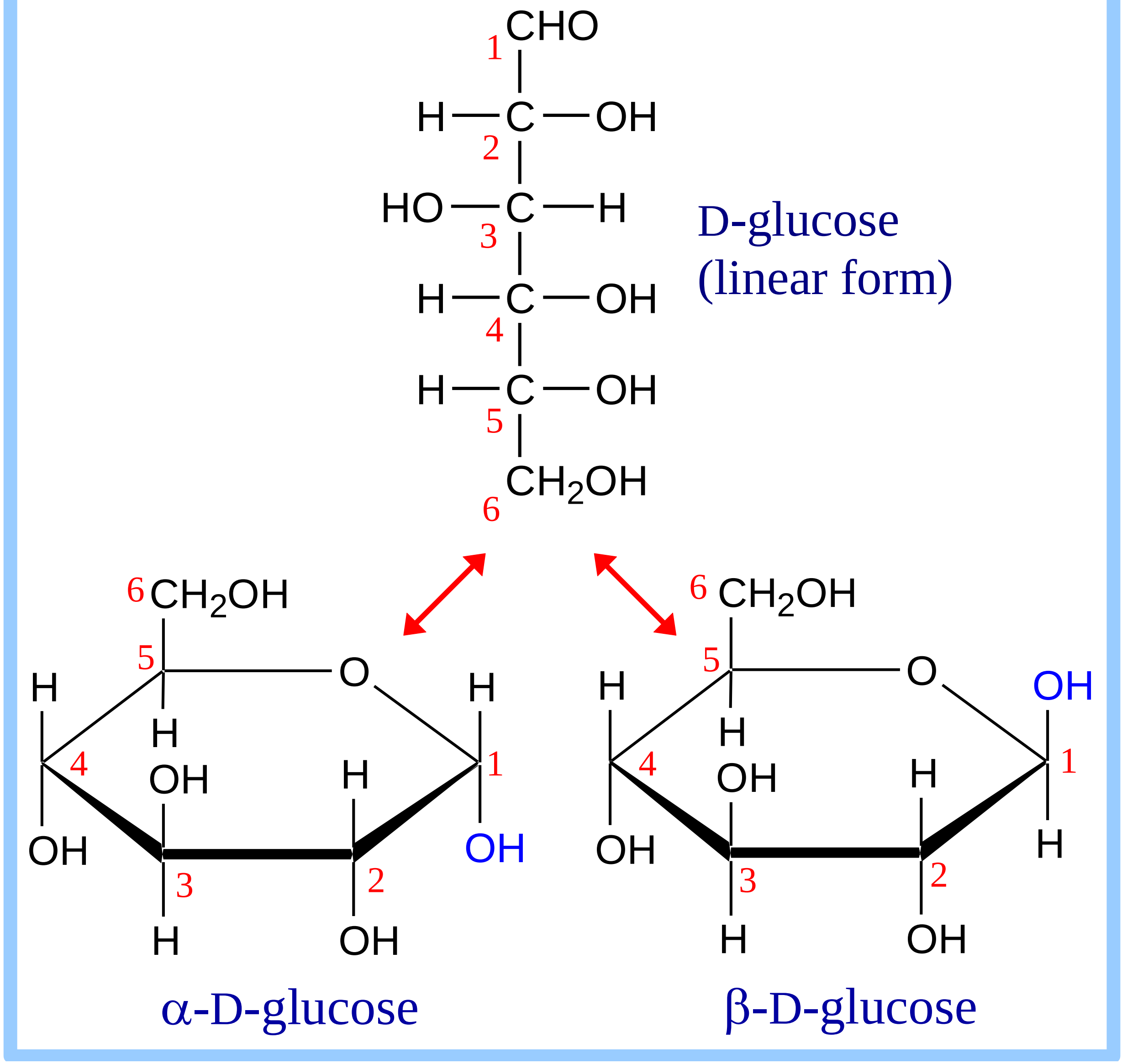
glucopyranose

new chiral center

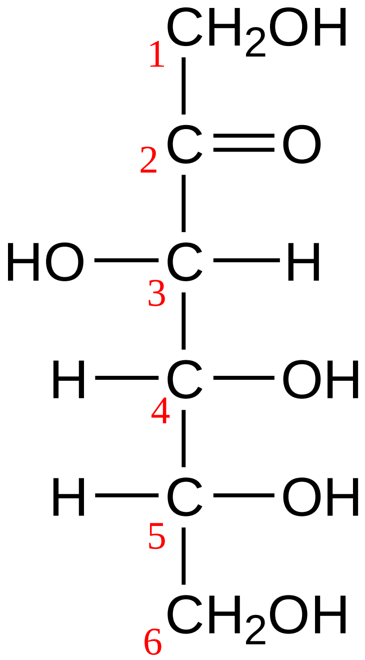
Note: the pyranose forms of carbohydrates adopt chair conformations.

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

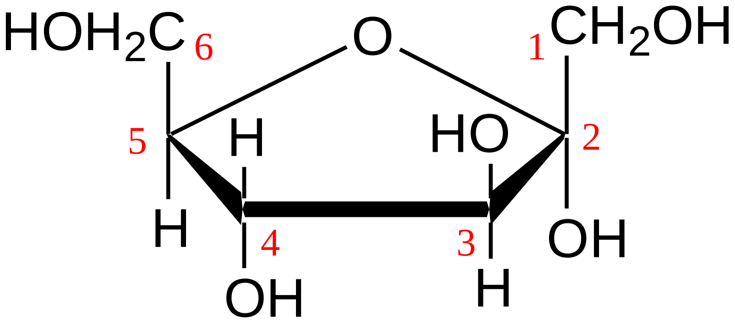
Glucose forms an intra-molecular hemiacetal, as the C1 aldehyde & C5 OH react, to form a 6-member pyranose ring, named after pyran.



These representations of the cyclic sugars are called Haworth projections.



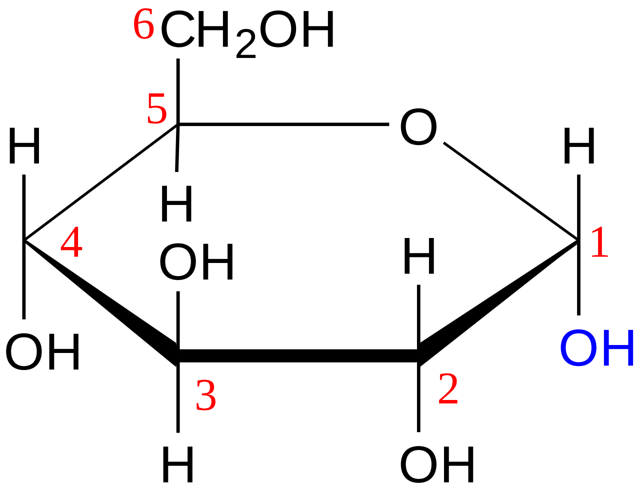
D-fructose (linear)



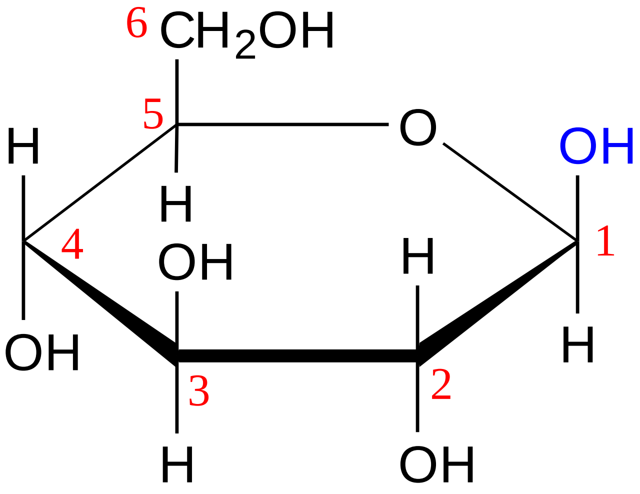
α-D-fructofuranose

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.



α-D-glucose

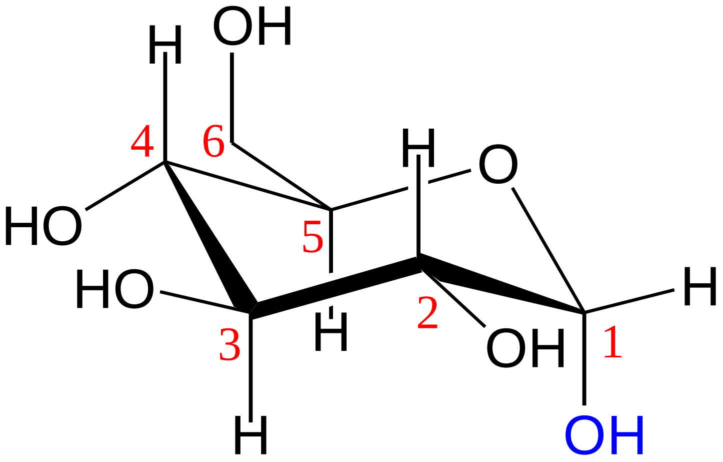


β-D-glucose

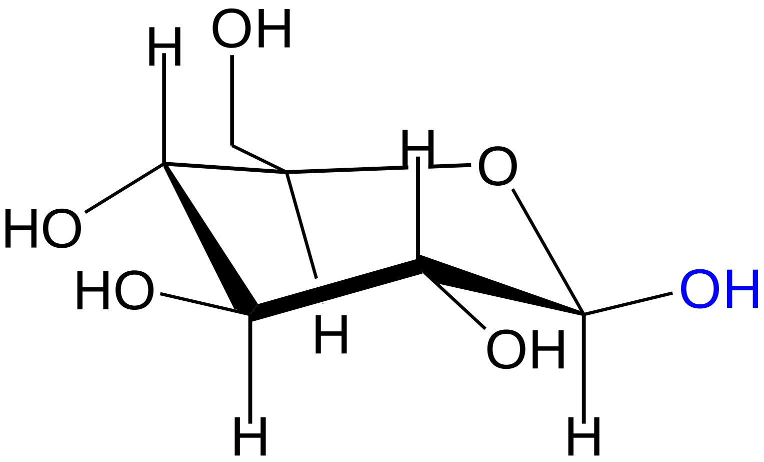
Cyclization of glucose produces a new asymmetric center at C1. The 2 stereoisomers are called anomers, α & β.

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

- α (OH below the ring)
- β (OH above the ring).



α-D-glucopyranose

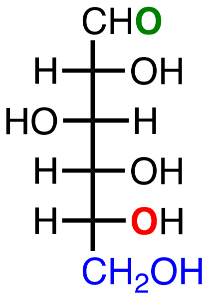


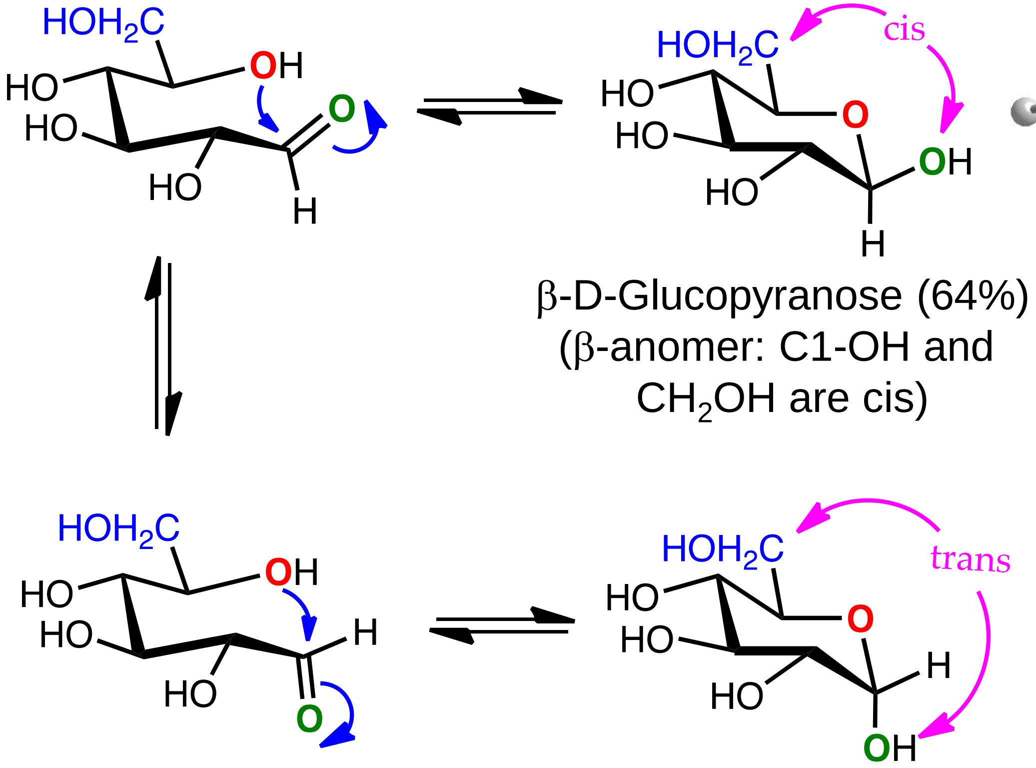
β-D-glucopyranose

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.

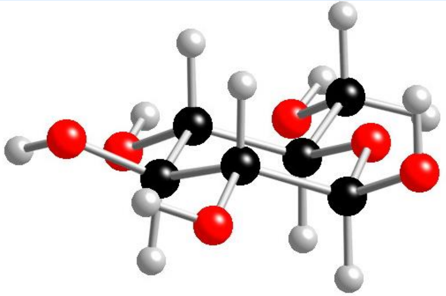
Mutarotation. The α- and β-anomers are in equilibrium, and interconvert through the open form. The pure anomers can be isolated by crystallization. When the pure anomers are dissolved in water they undergo mutarotation, the process by which they return to an equilibrium mixture of the anomer.



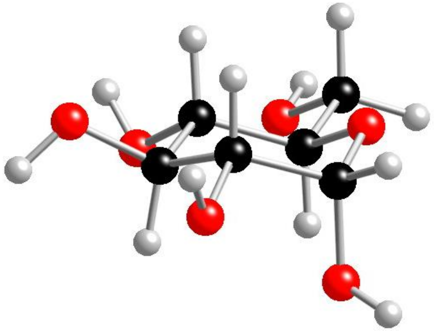


β-D-Glucopyranose (64%)
(β-anomer: C1-OH and CH_2OH are cis)

α-D-Glucopyranose (36%)
(α-anomer: C1-OH and CH_2OH are trans)



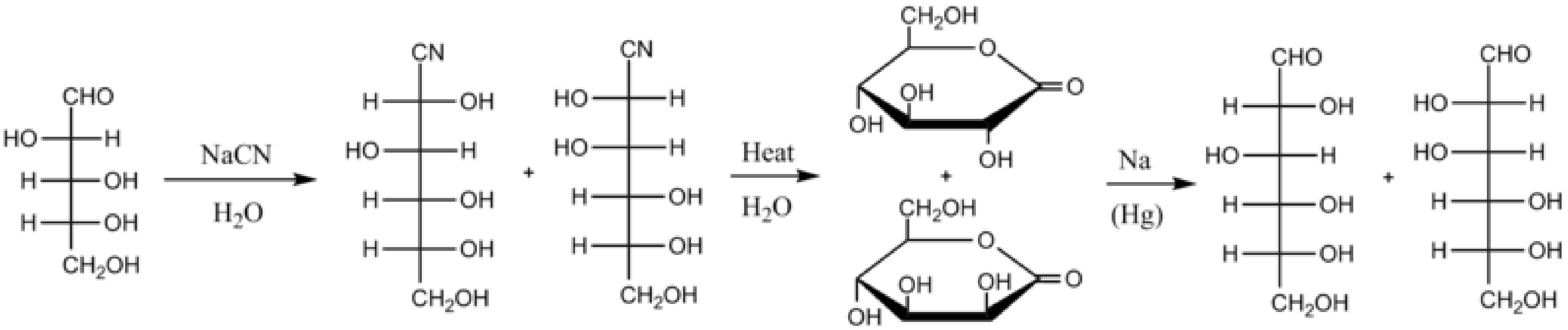
$[\alpha]_D$ +18.7°



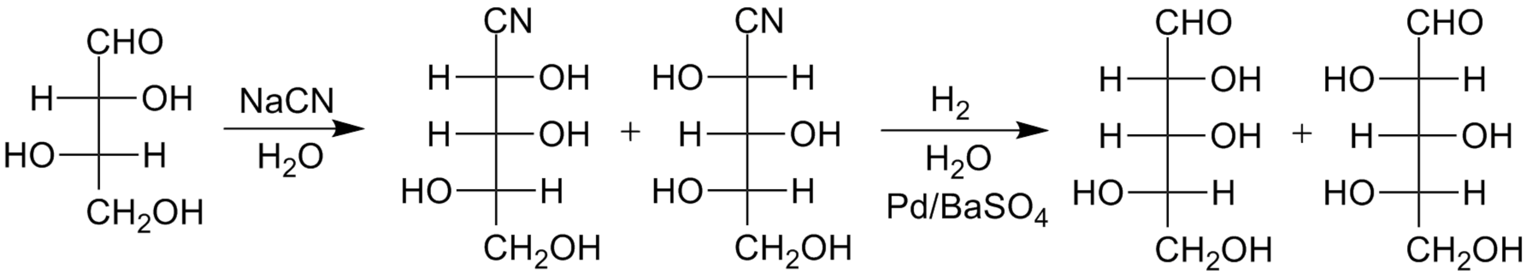
$[\alpha]_D$ +112.2°

Kiliani–Fischer synthesis

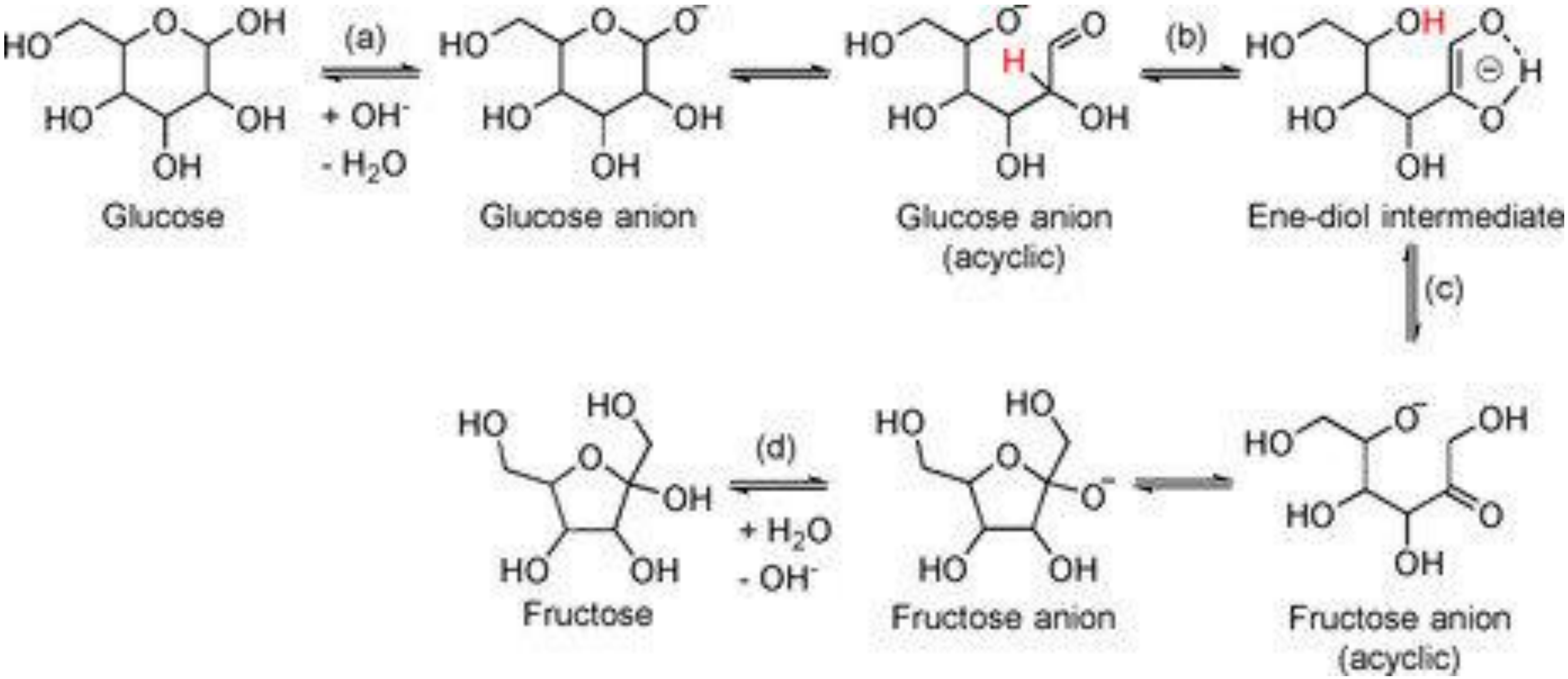
A series of reactions that extends the carbon chain in a carbohydrate by one carbon and one chiral center. e.g. conversion of aldopentose into an aldohexose.



Pd/BaSO4 in presence of hydrogen gas is improved reagent used in Kiliani-Fisher synthesis as shown below.



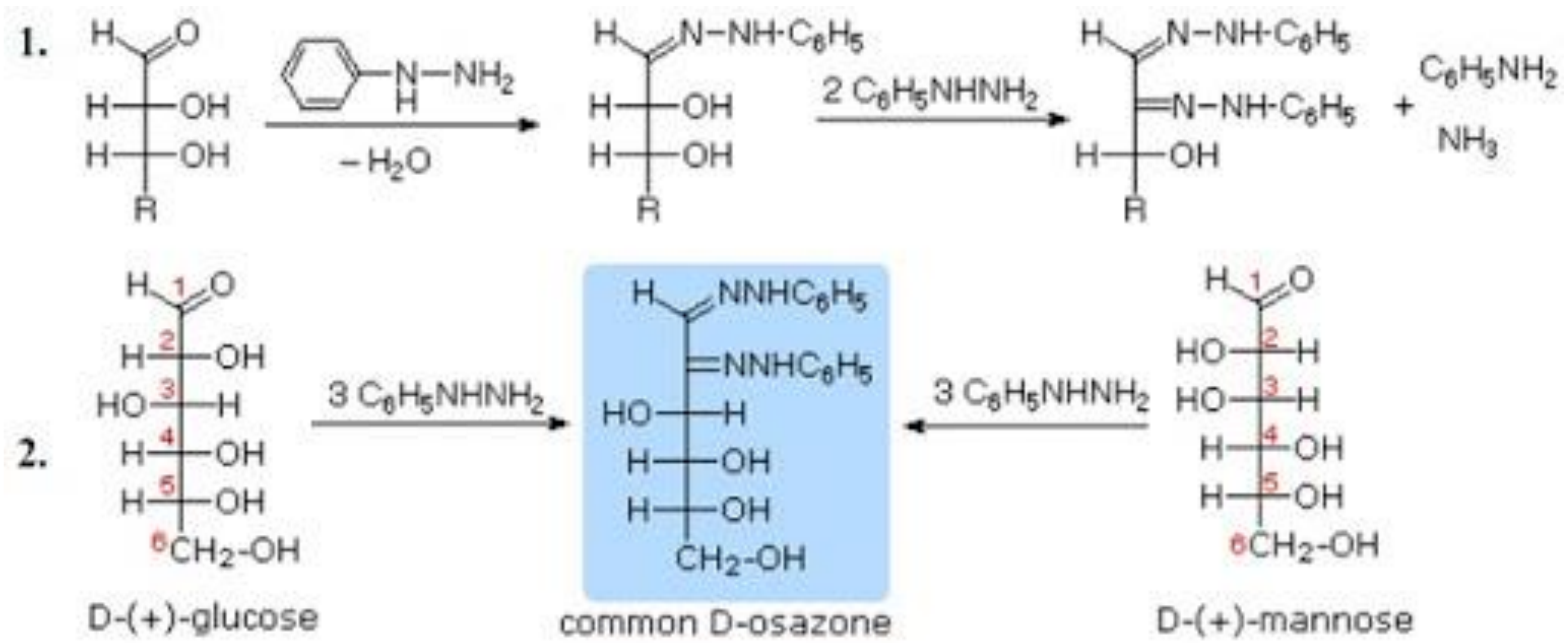
Conversion of Aldoses to ketoses



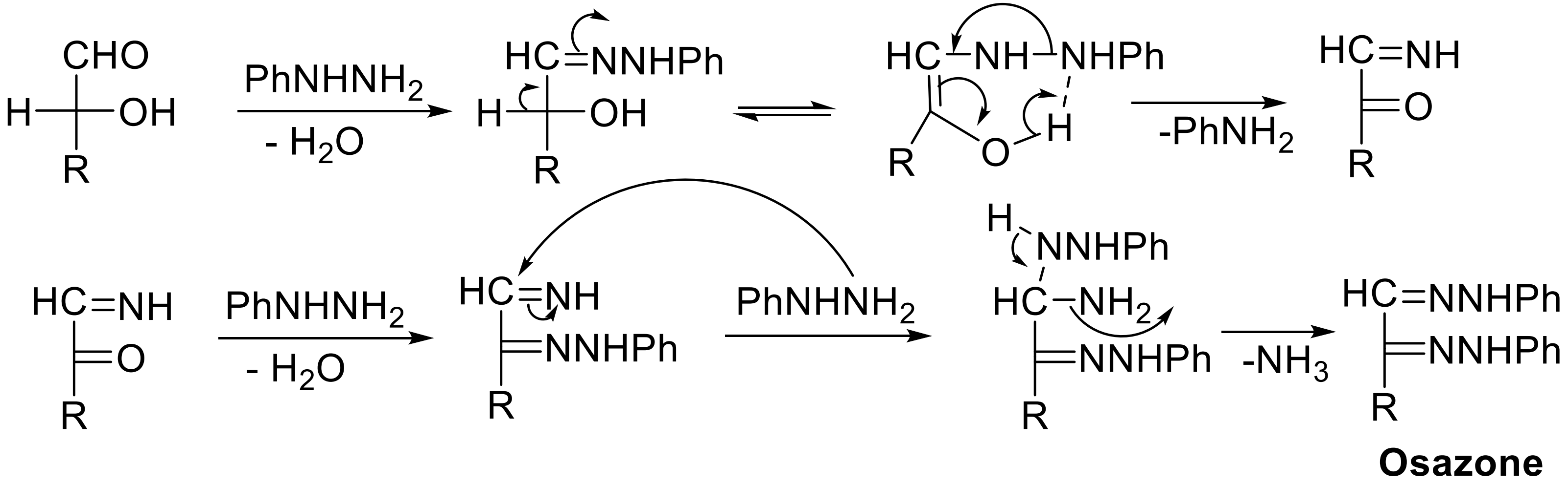
Reproduced from- Catal. Sci. Technol., 2017, 7, 5321–5331, ACS Catal., 2015, 5, 3162

Osazone formation

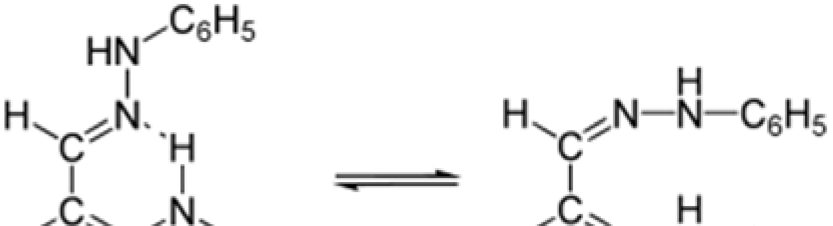
When aldoses and ketoses reacts with excess of phenylhydrazine they forms a common product called as osazone. Aniline and ammonia forms as byproduct in this reaction. The overall reaction is shown below.



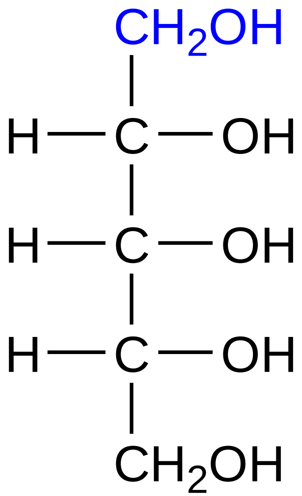
Mechanism of Osazone formation



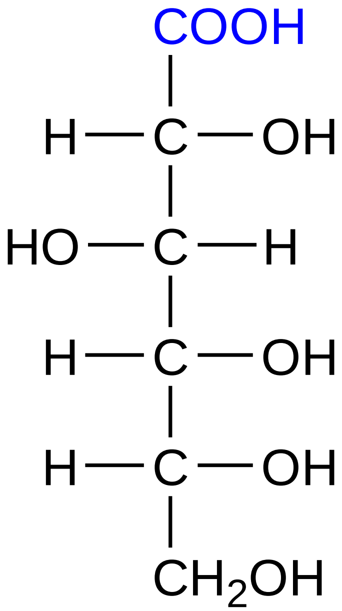
Similar reaction mechanism was followed in case of ketoses. In case of ketoses the first reaction of phenylhydrazine took place at C-2 instead C-1. The Osazone is stable compound, it forms six membered stable structure as shown below. Therefore further reaction (at C-3) of phenylhydrazine on osazone didn’t took place.



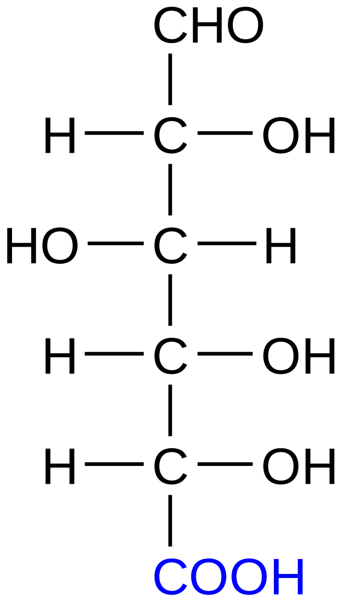
Sugar derivatives



D-ribitol



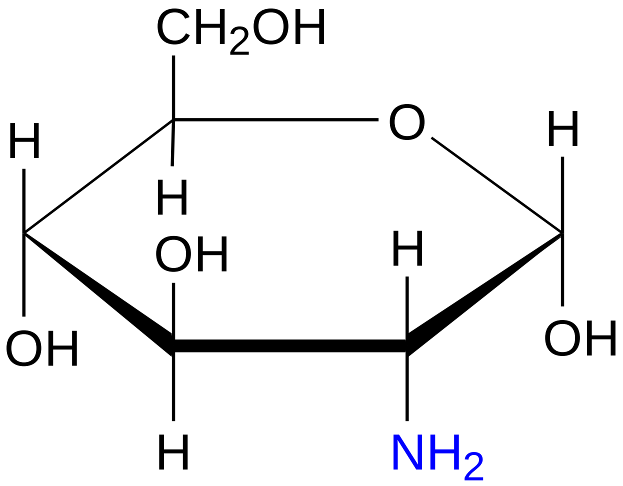
D-gluconic acid



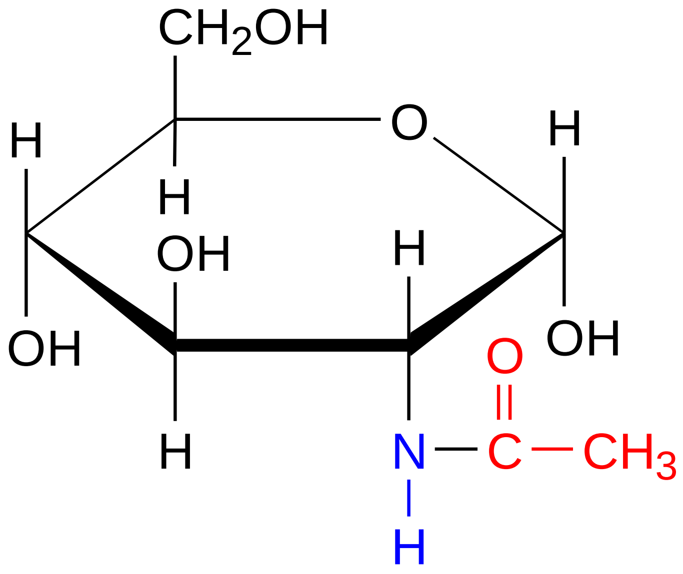
D-glucuronic acid

- 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



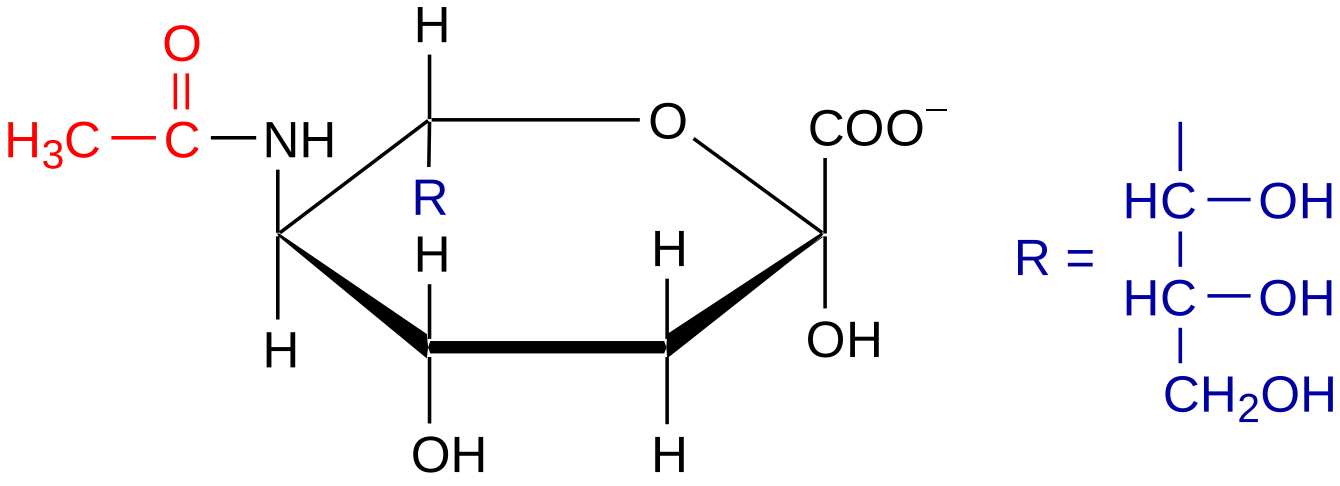
α-D-glucosamine



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.

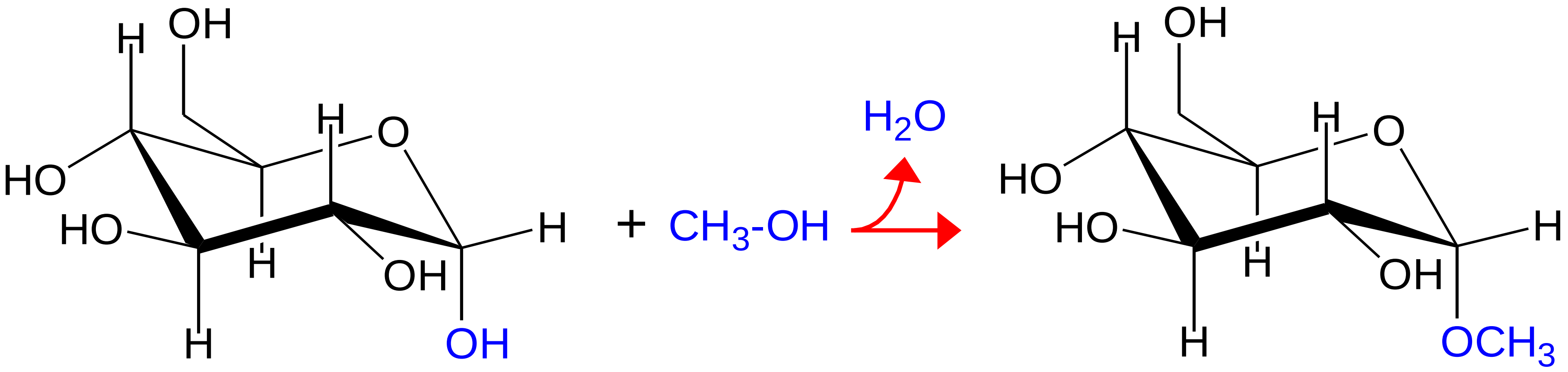
Sialic acid imparts negative charge to glycoproteins, because its carboxyl group tends to dissociate a proton at physiological pH, as shown here.

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:

$$\mathrm{R\text{-}OH} + \mathrm{HO\text{-}R'} \rightarrow \mathrm{R\text{-}O\text{-}R'} + \mathrm{H_2O}$$

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

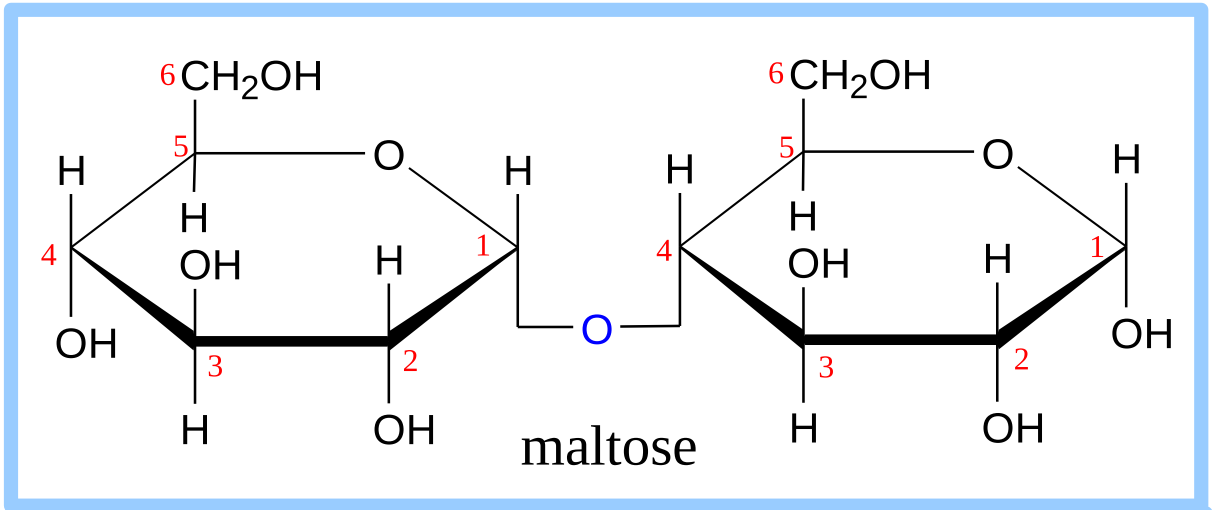


α-D-glucopyranose methanol methyl-α-D-glucopyranose

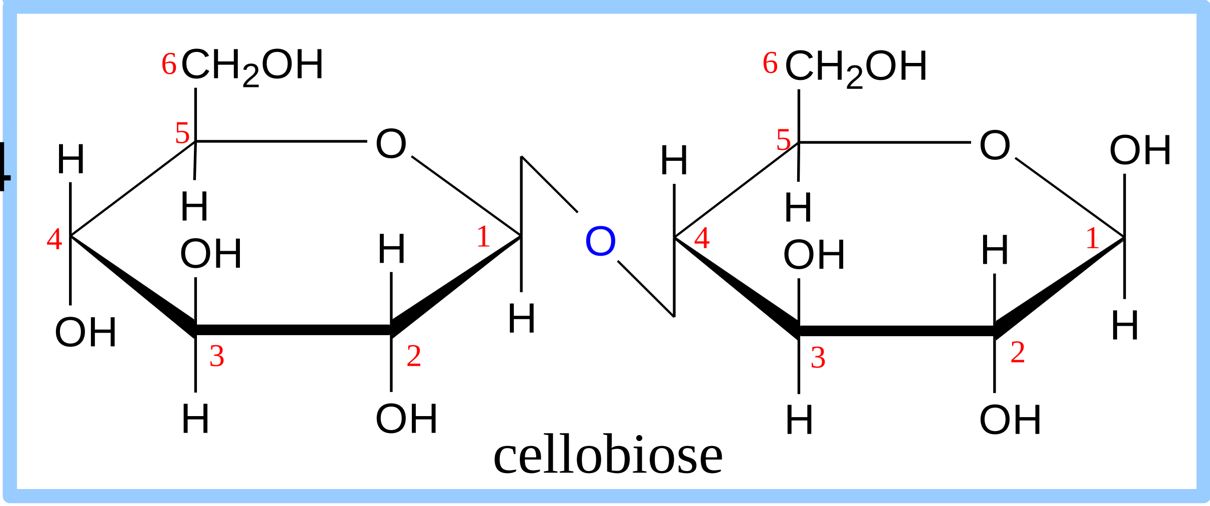
Disaccharides:

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 **β** anomer (O on C1 points up).

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

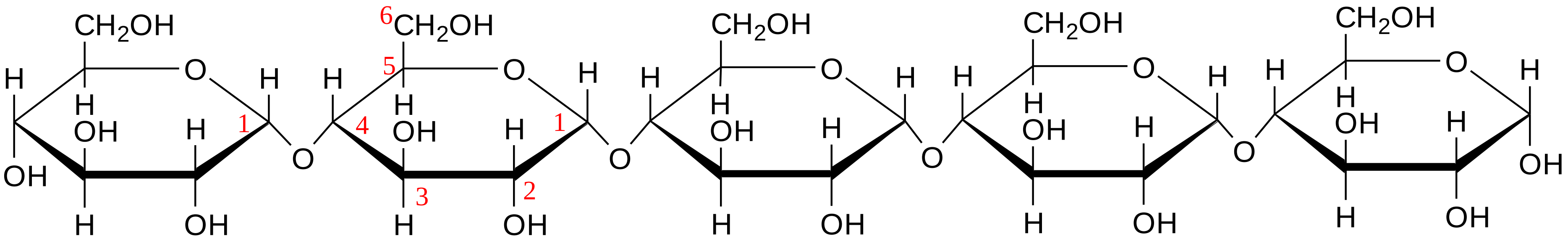
Other **disaccharides** include:

- **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 α (O points down from ring), the linkage is α(1→2).

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

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



amylose

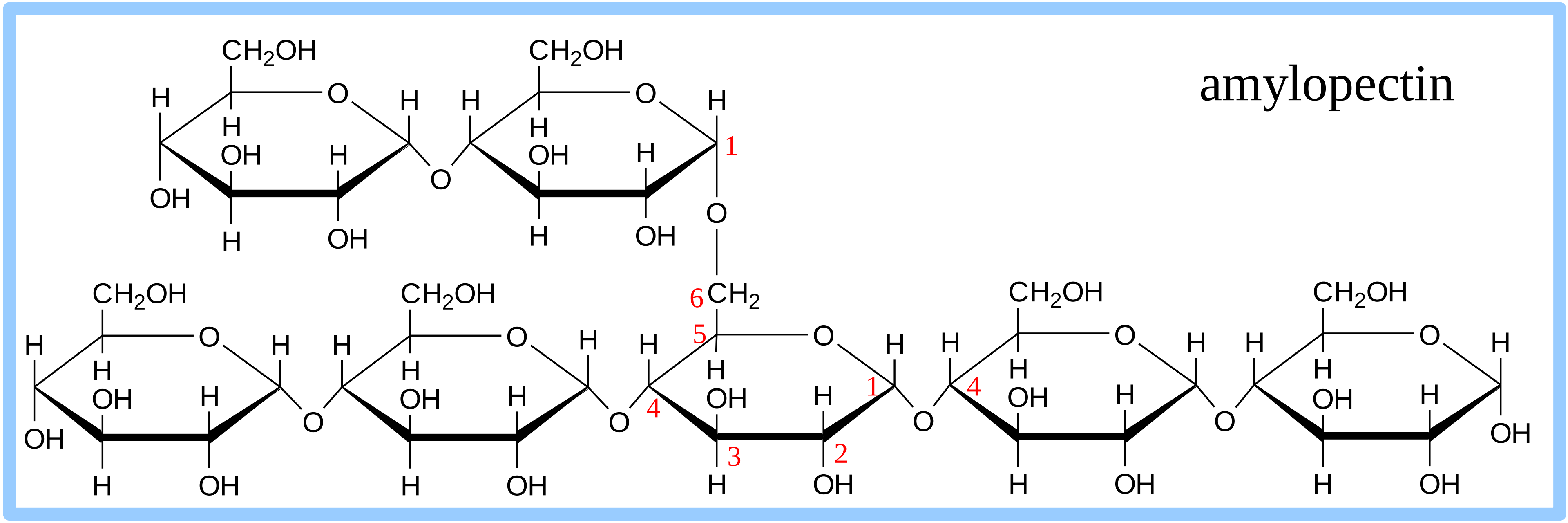
Polysaccharides:

Plants store glucose as **amylose** or **amylopectin**, glucose polymers collectively called starch.

Glucose storage in **polymeric** form **minimizes osmotic effects**.

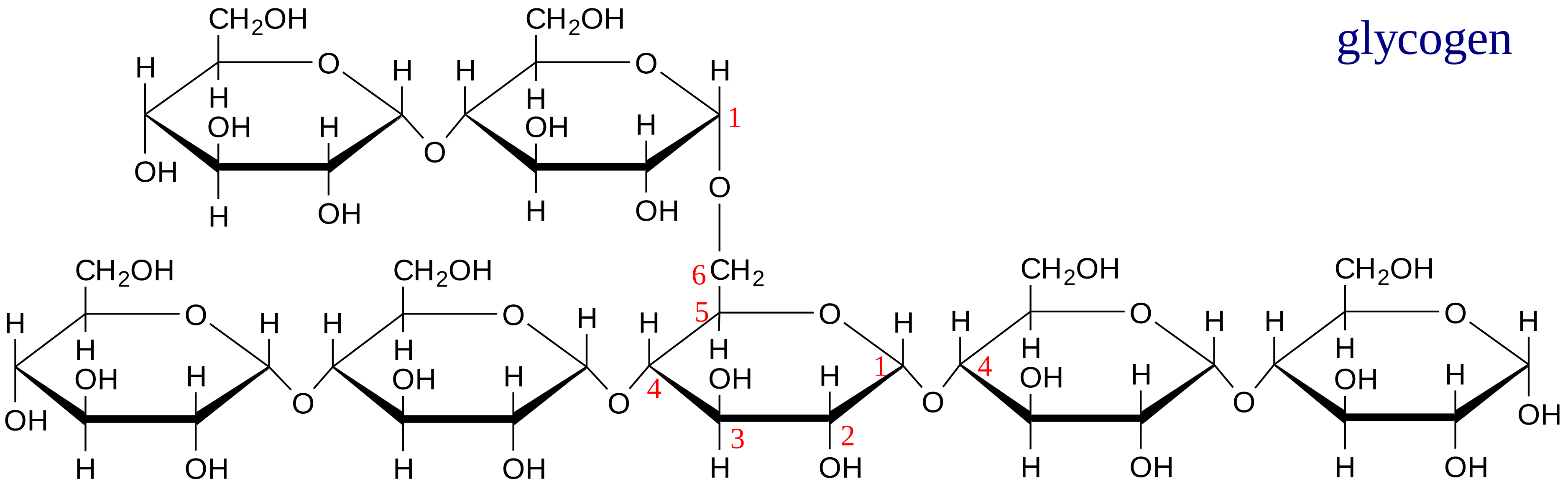
Amylose is a glucose polymer with **α(1→4)** linkages.

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 α(1→4) linkages, but it also has **branches** formed by α(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.



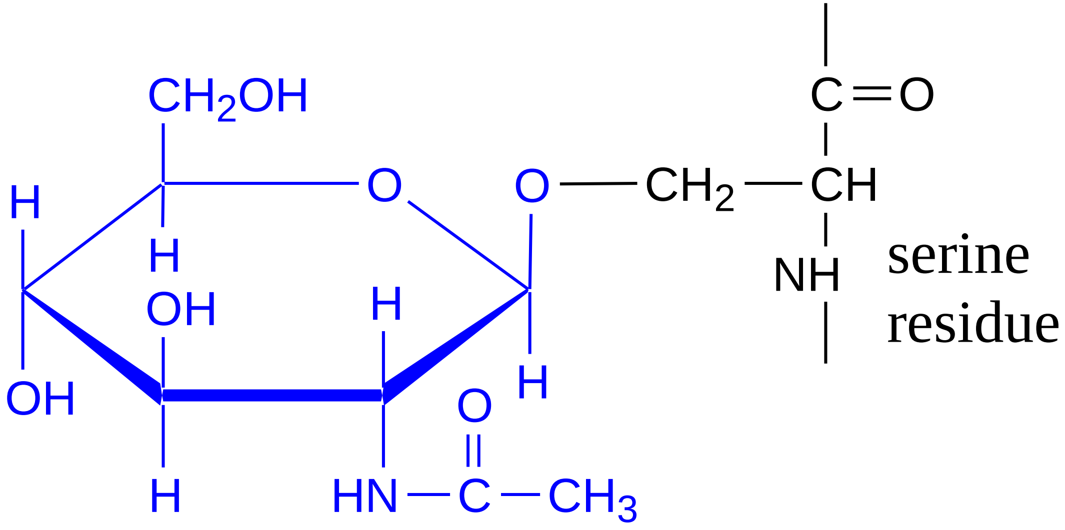
Glycogen, the glucose storage polymer in **animals**, is similar in structure to amylopectin.

But glycogen has **more** α(1→6) **branches**.

The highly branched structure permits rapid glucose release from glycogen stores, e.g., in muscle during exercise.

The ability to rapidly mobilize glucose is more essential to animals than to plants.

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

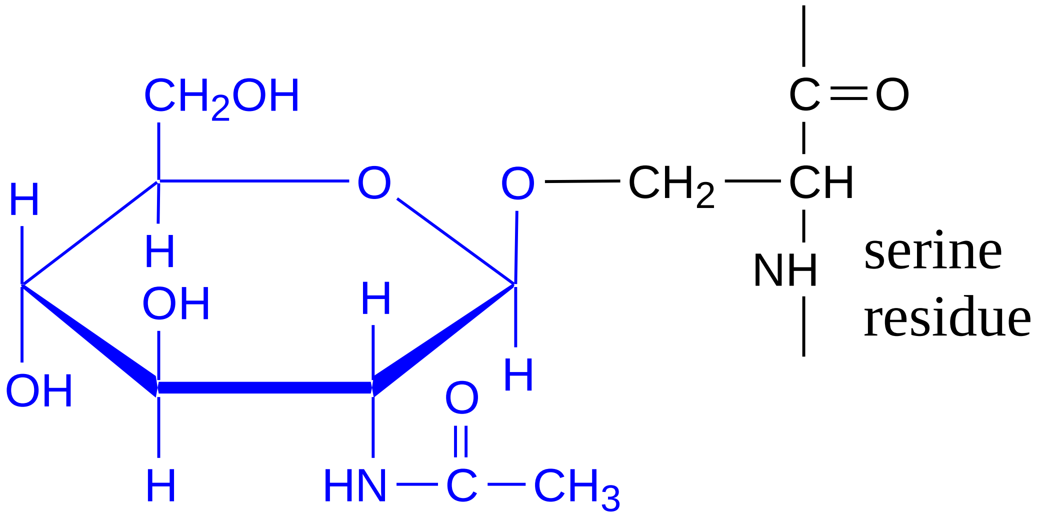


β-D-*N*-acetylglucosamine

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**.

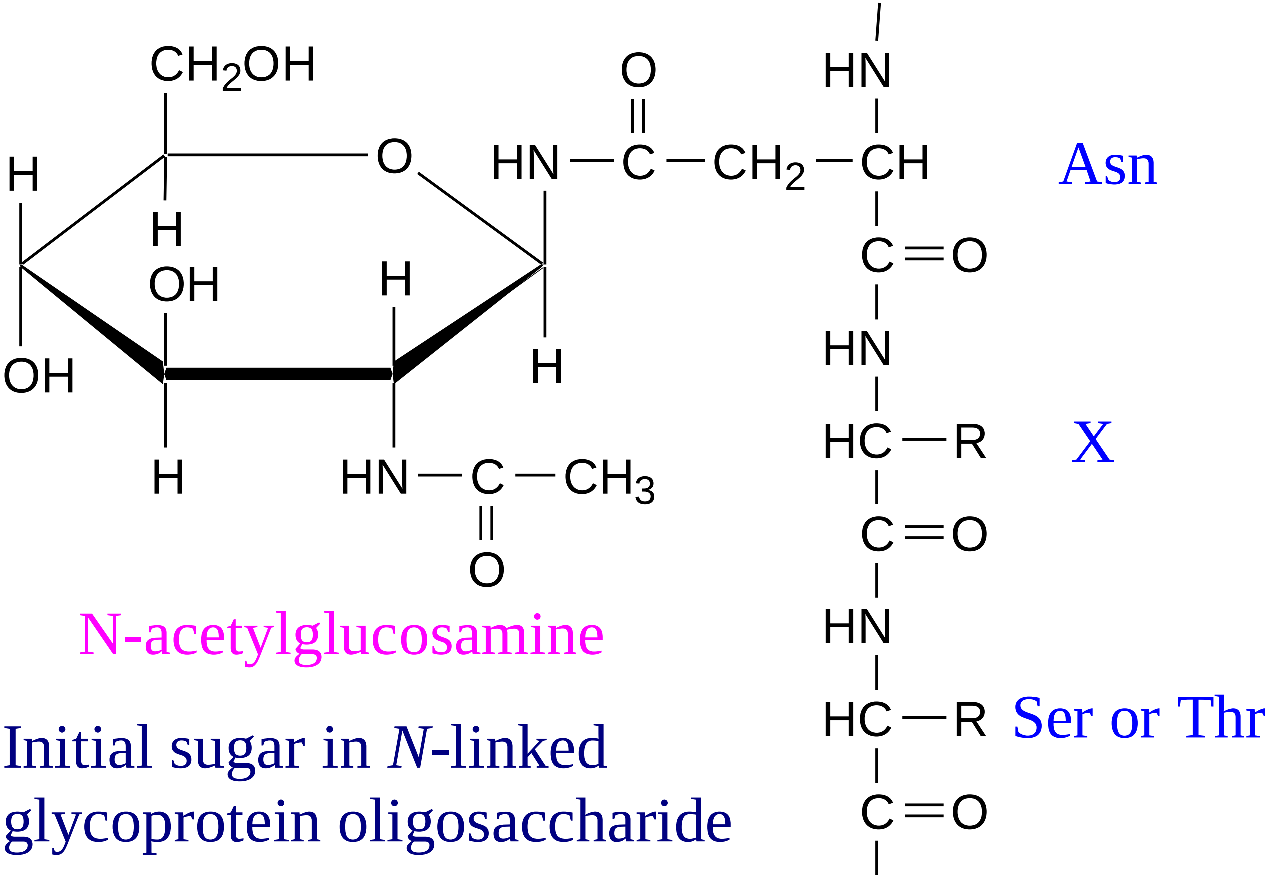


β-D-*N*-acetylglucosamine

***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).

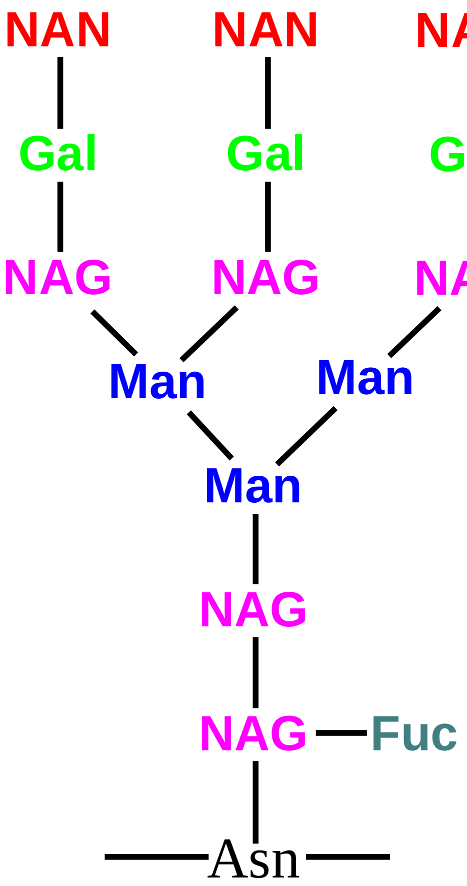


N-acetylglucosamine

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.


NAN NAN NAN
Gal Gal Gal
NAG NAG NAG
Man Man
Man
NAG
NAG — Fuc
— Asn —

***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**.

Concanavalin A & **wheat germ agglutinin** are plant lectins that have been useful research tools.

The **C-type lectin-like domain** is a **Ca^{++}-binding** carbohydrate recognition domain in many **animal lectins**.

Recognition/binding of CHO moieties of glycoproteins, glycolipids & proteoglycans **by** animal **lectins** is a factor in:

- cell-cell recognition
- adhesion of cells to the extracellular matrix
- interaction of cells with chemokines and growth factors
- recognition of disease-causing microorganisms
- initiation and control of inflammation.