Fundamentals of Glycan Structure

Learning Objectives

- ***How are glycans named?***
- ***What are the different constituents of a glycan?***
- ***How are these represented?***
- ***What conformations do sugar residues adopt in solution***
- ***Why do glycan conformations matter?***

Fundamentals of Glycan Structure

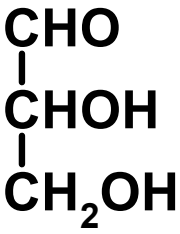
- ***Carbohydrate Nomenclature***
- ***Monosaccharides***
  - ✓ *Structure*
  - ✓ *Fisher Representation*
  - ✓ *Cyclic Form*
  - ✓ *Chair Form*
  - ✓ *Mutarotation*
- ***Monosaccharide Derivatives***
  - ✓ *Reducing Sugars*
  - ✓ *Uronic Acids*
  - ✓ *Other Derivatives*
- ***Monosaccharide Conformation***
- ***Inter-Glycosidic Bond***
  - ✓ *Normal*
  - ✓ *Sucrose*
  - ✓ *Lactose*
  - ✓ *Sequence Specificity and Recognition*
  - ✓ *Branching*

Carbohydrate Nomenclature

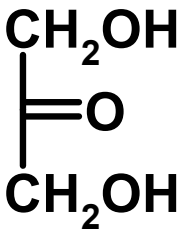
- ***The word 'carbohydrate' implies "hydrate of carbon" … $C_n(H_2O)_m$***
  - ✓ *Glucose (a monosaccharide) $C_6H_{12}O_6$ … $C_6(H_2O)_6$*
  - ✓ *Sucrose (a disaccharide) $C_{12}H_{22}O_{11}$ … $C_{12}(H_2O)_{11}$*
  - ✓ *Cellulose (a polysaccharide) $(C_6H_{12}O_6)_n$… $(C_6(H_2O)_6)_n$*
- ***Not all carbohydrates have this formula … some have nitrogen***
  - ✓ *Glucosamine (glucose + amine) …. $C_6H_{13}O_5N$ … $-NH_2$ at the 2-position of glucose*
  - ✓ *N-acetyl galactosamine (galactose + amine + acetyl group) …. $C_8H_{15}O_6N$ … $-NHCOCH_3$ at the 2-position of galactose*
- ***Typical prefixes and suffixes used in naming carbohydrates***
  - ✓ *Suffix = '**-ose**' & prefix = '**tri-**', '**tetr-**', '**pent-**', '**hex-**'*
  - ✓ *Pentose (a five carbon monosaccharide) or hexose (a six carbon monosaccharide)*
- ***Functional group types***
  - ✓ *Monosaccharides with an aldehyde group are called aldoses … e.g., glyceraldehyde*
  - ✓ *Those with a keto group are called ketoses … e.g., dihydroxyacetone*

Monosaccharides

- ***Structure***
  - ✓ *Have a general formula $C_nH_{2n}O_n$ and contain a carbonyl group*
  - ✓ *Common monosaccharides have 3 → 9 carbons*
  - ✓ *Two molecules form the class of smallest monosaccharides … glyceraldehyde and dihydroxyacetone*



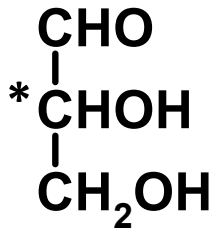
Glyceraldehyde



Dihydroxyacetone

Monosaccharides

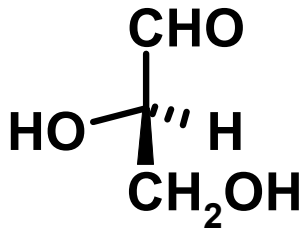
- ***Glyceraldehyde has a chiral carbon***



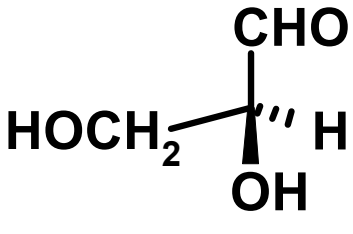
* = chiral carbon

… 2 stereoisomers

- ***3-Dimensional arrangement***

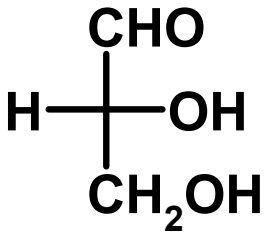


(R)

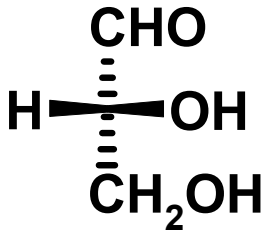


(S)

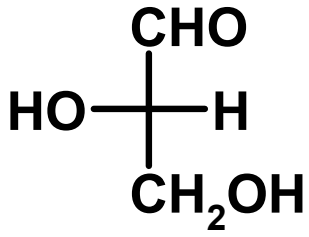
- ***2-Dimensional representation***



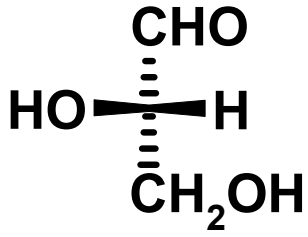
(D)



(R)



(L)

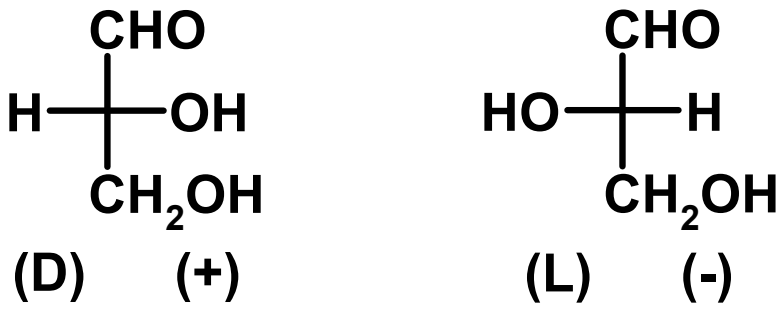


(S)

Monosaccharides

- ***Emil Fischer's representation***

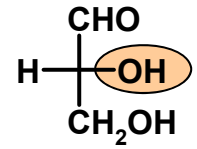
  - ✓ *Arbitrarily assigned the dextrorotatory enantiomer as D-glyceraldehyde, which fortuitously proved correct*



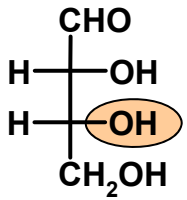
***Note**: This does not mean that all D-saccharides are dextrorotatory; likewise it does not mean that all D-saccharides are also (R) in the Cahn-Ingold-Prelog system*

Monosaccharides

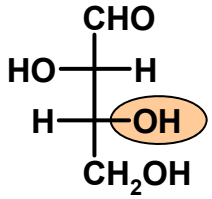
- ***'D' monosaccharides***



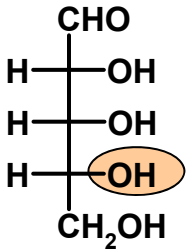
D-glyceraldehye



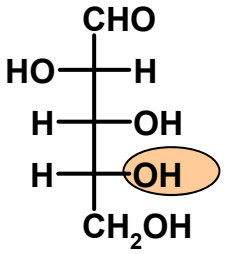
D-erythrose



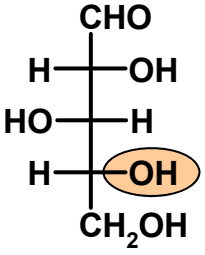
D-threose



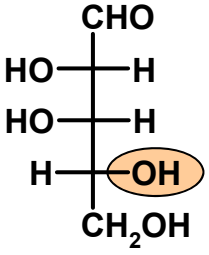
D-ribose



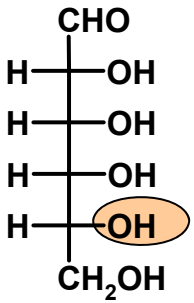
D-arabinose



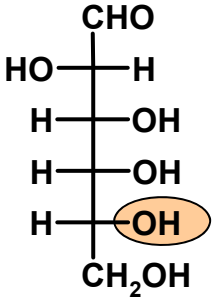
D-xylose



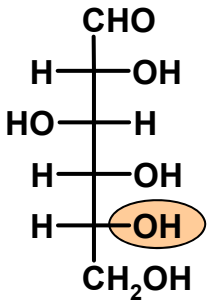
D-lyxose



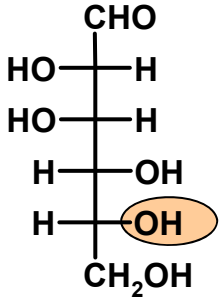
D-allose



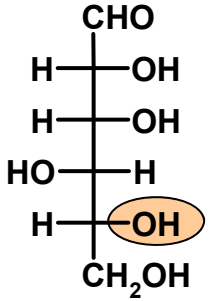
D-altrose



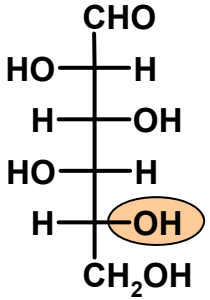
D-glucose



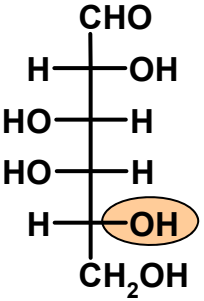
D-mannose



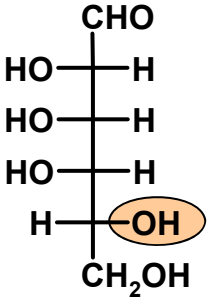
D-gulose



D-idose



D-galactose

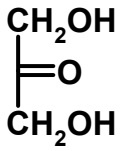


D-talose

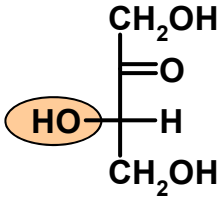
- ✓ *'D' monosaccharides have the same configuration on their penultimate carbon as 'D' glyceraldehyde*
- ✓ *Exactly similar series for 'L' monosaccharides starting with 'L' glyceraldehyde*
- ✓ *Likewise, similar series exists for 'L' and 'D' ketoses starting from dihydroxyacetone*

Monosaccharides

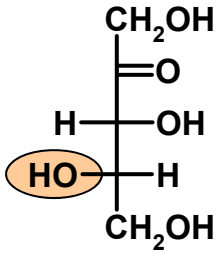
- ***'L' ketoses***



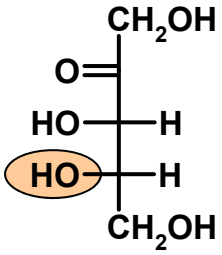
Dihydroxyacetone
(no D- or L-form)



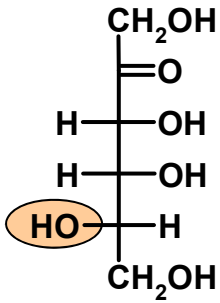
L-erythrulose



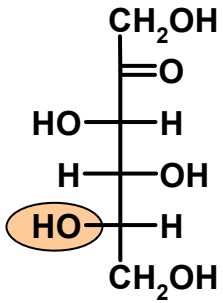
L-xylulose



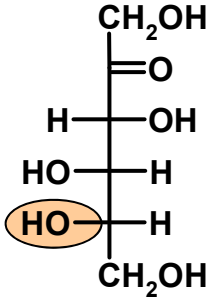
L-ribulose



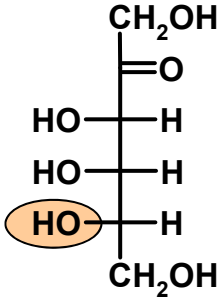
L-psicose



L-fructose



L-sorbose



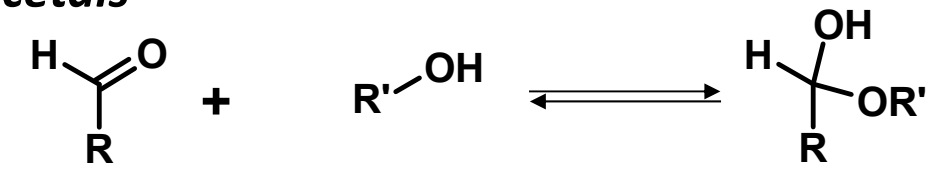
L-tagatose

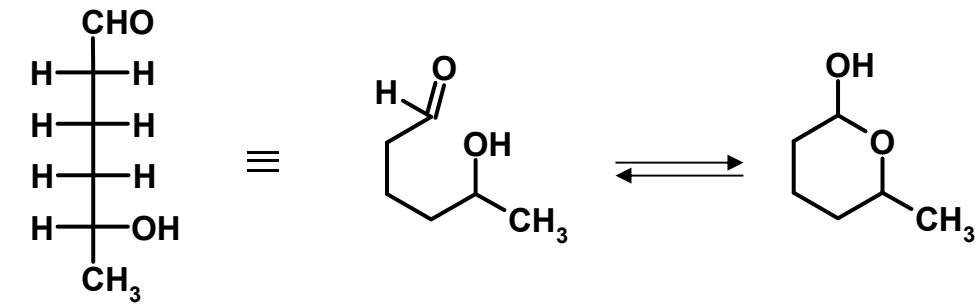
Monosaccharides

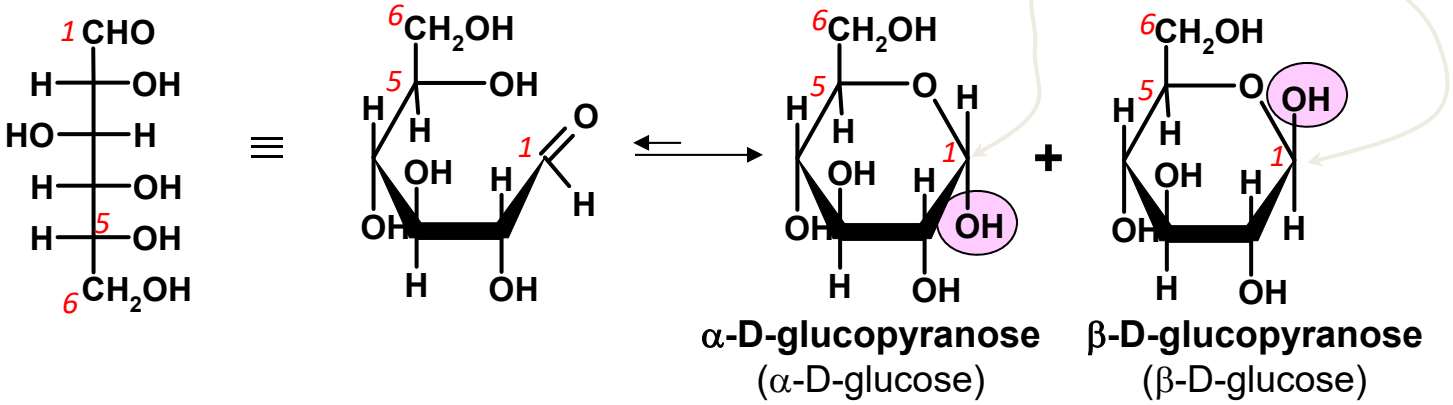
- ***Some Useful Information on Monosaccharides***
  - ✓ *D-glucose is also called dextrose because it is dextrorotatory.*
  - ✓ *D-glucose is the most abundant monosaccharide; it is present in most fruit juices*
  - ✓ *D-glucose is also called blood sugar …… it is present in about 65 – 100 mg per 100 mL of blood*
  - ✓ *The dextrose solution for iv drip contains about 5% glucose + 0.15% saline (KCl)*
  - ✓ *Fructose is one of the monosaccharides present in disaccharide sucrose (table sugar) … the other is D-glucose*
  - ✓ *Galactose is part of lactose (milk sugar)*

Monosaccharides

- ***Monosaccharides Are Typically not Linear (or Open-Chain)***
  - ✓ *Remember from organic chemistry: Aldehydes and ketones react with alcohols to give* ***hemi-acetals***



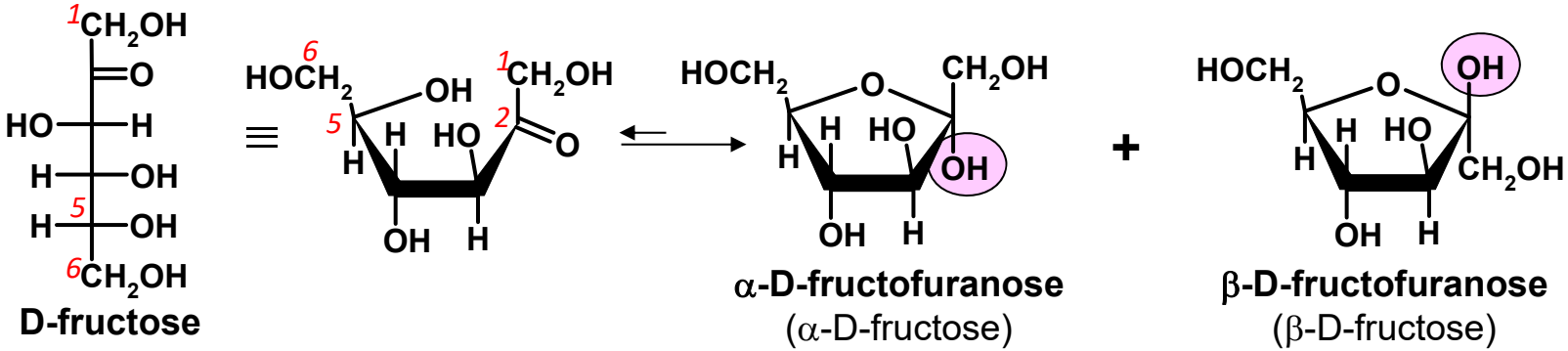




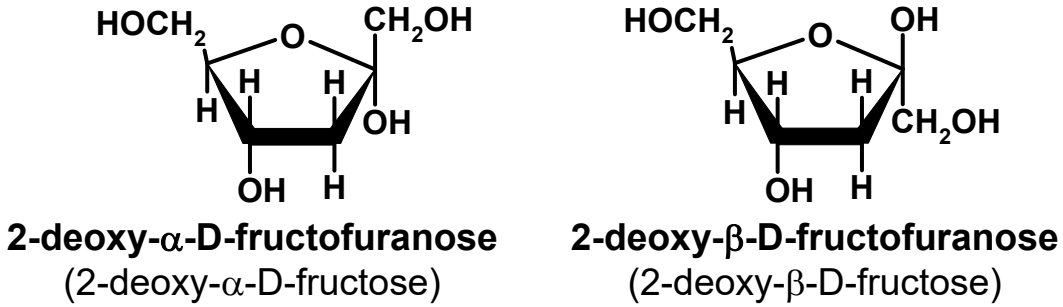
- ✓ *What is an anomer?*
- ✓ *How are α- and β-anomers defined?*
- ✓ *What is an epimer?*

Monosaccharides

- ***Cyclic Structure of Ketoses***
  - ✓ *Similar cyclic structures exist for five membered saccharides … e.g., ribose.*
  - ✓ *These five membered cyclic hemiacetals are well known for nucleic acids … e.g., RNA (**ribo**nucleic acid) and DNA (**deoxyribo**nucleic acid)*

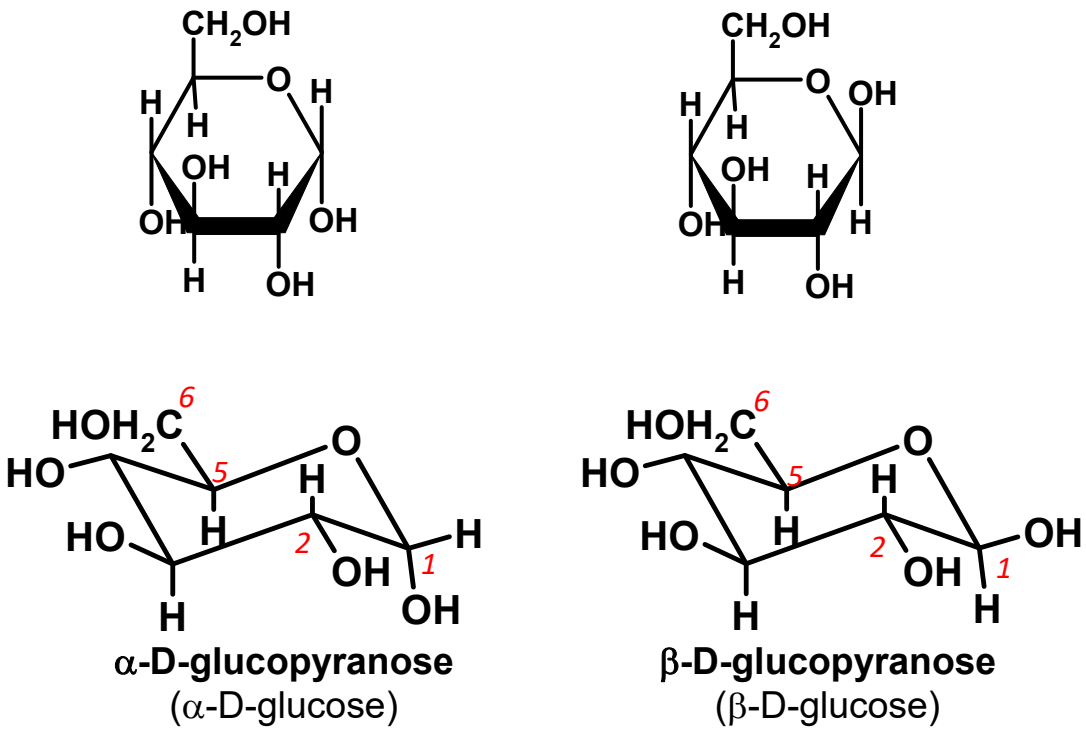


HAWORTH Projections



Monosaccharides

- ***The Cyclic Form Assumes Different Conformations in Solution***

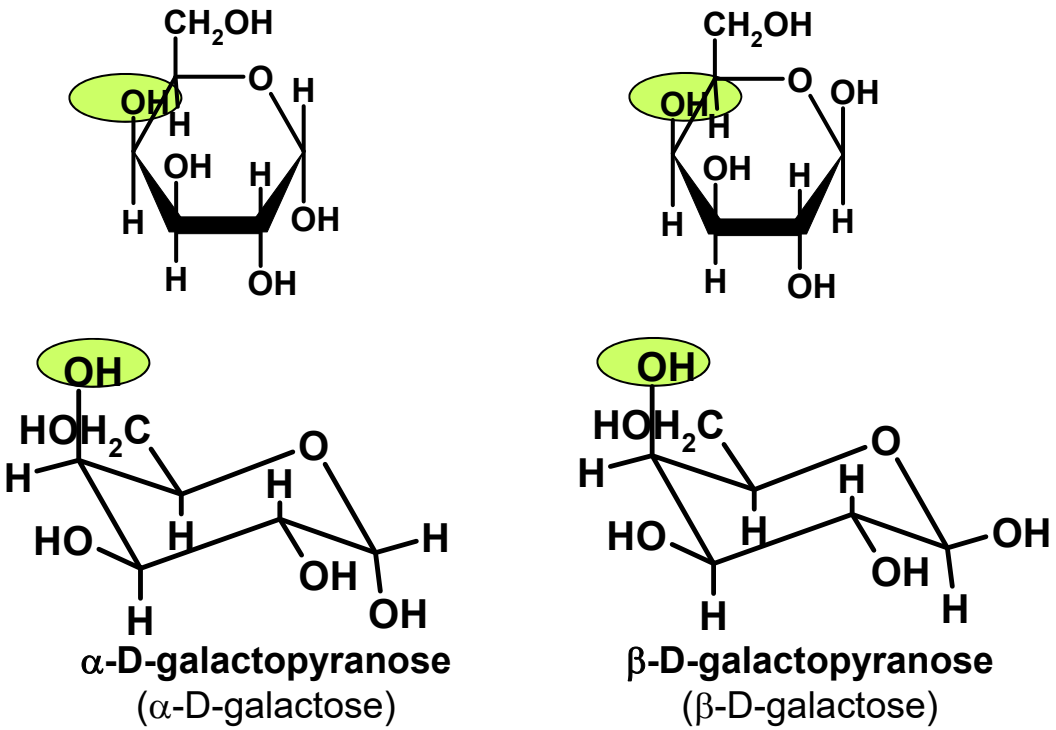


Chair Conformations

- ✓ *Why is β-D-glucose the most common sugar in nature?*

Monosaccharides

- ***The Cyclic Form of D-Galactose***



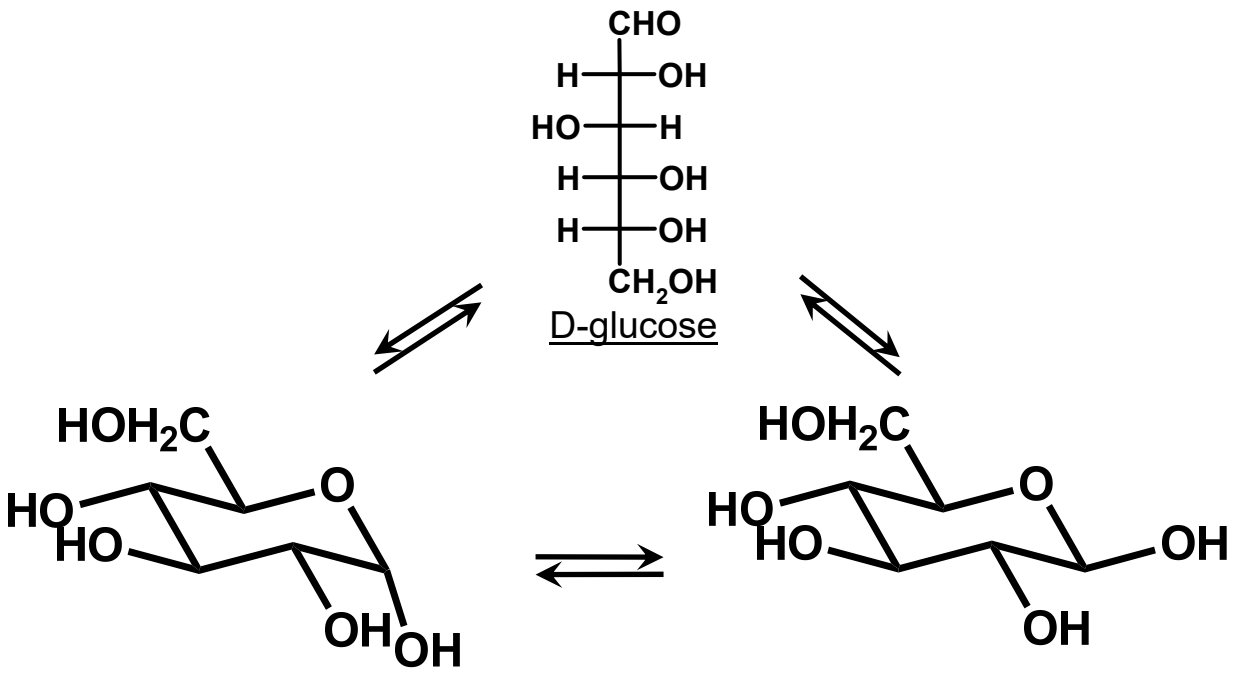
α-D-galactopyranose
(α-D-galactose)

β-D-galactopyranose
(β-D-galactose)

Chair Conformations

Monosaccharides

- ***Mutarotation Arises From Cyclic and Open-Chain Form Interconversion***



α-D-glucopyranose
(α-D-glucose)

β-D-glucopyranose
(β-D-glucose)

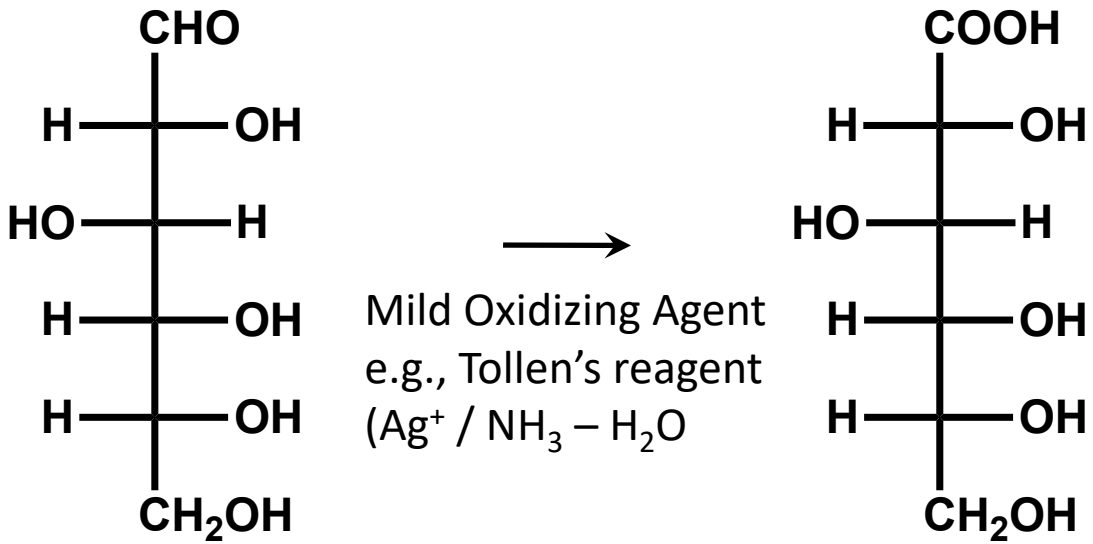
Optical activity

$[\alpha]_D = +112.2^O$ **$[\alpha]_D = +18.7^O$**

At equilibrium …. $[\alpha]_D = +52.7^O$ (Not an average!
But weighted average; 63.6% of the β-anomer)

Monosaccharide Derivatives

- ***Reducing Sugars and Reducing End***

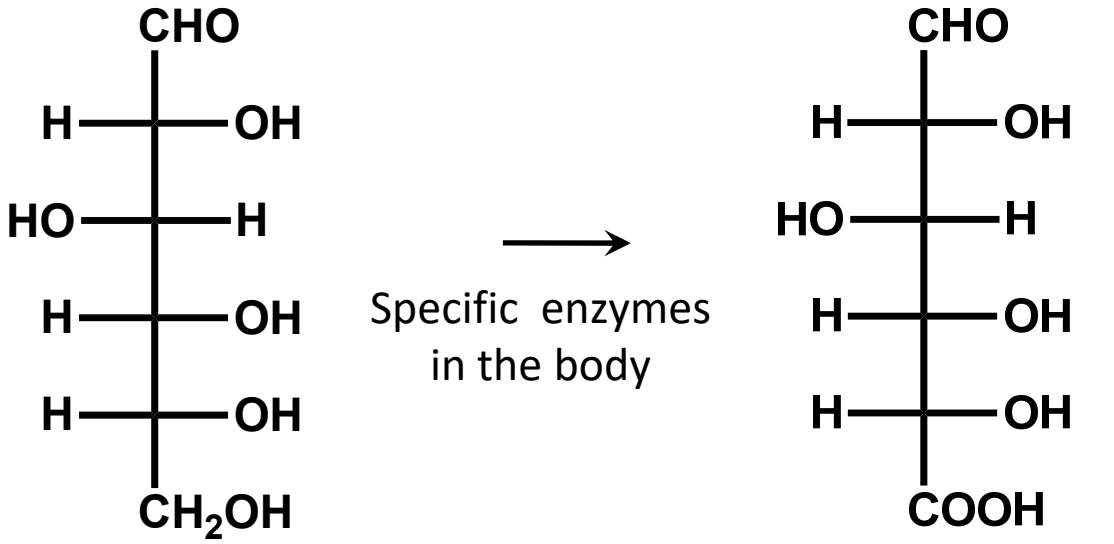


D-Glucose **D-Gluconic Acid**

- ✓ *Reducing sugars and diabetes*
  - Hemoglobin (Hb) is glycated ….. HbA1c levels should be < 6.5%
  - Reaction of high glucose with NH_2 group at the N-terminus (Val) of Hb
  - Non-enzymatic
  - Reflects Glc levels for the past 3 months (typical life of erythrocytes)

Monosaccharide Derivatives

- ***Uronic Acids***

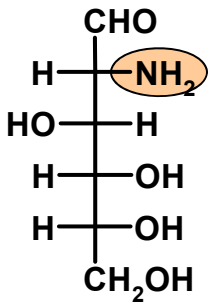


D-Glucose **D-Glucuronic Acid**

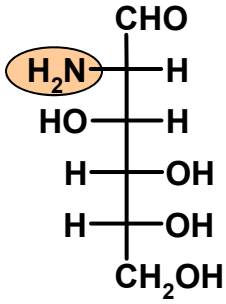
- ✓ *Biological relevance*
  - Several other uronic acids occur in our body, e.g., D-mannuronic acid (from D-mannose), L-iduronic acid (from L-idose), D-galacturonic acid
  - Components of many polysaccharides
  - Add to the conformational complexity , e.g., L-iduronic acid can exist in several conformations other than chair – $^{2}S_{O}$, $^{O}S_{2}$, $^{1}C_{4}$ and $^{4}C_{1}$

Monosaccharide Derivatives

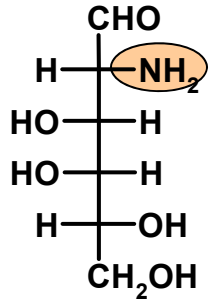
- ***Glycosamines***
  - ✓ *Several aminosugars are known including D-glucosamine, D-mannosamine, D-galactosamine and N-acetyl-D-glucosamine. These have a nitrogen at position 2*



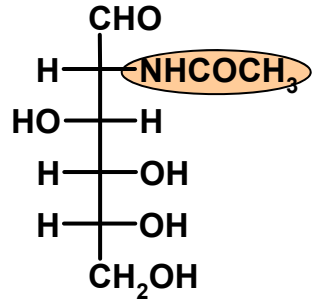
D-glucosamine



D-mannosamine



D-galactosamine

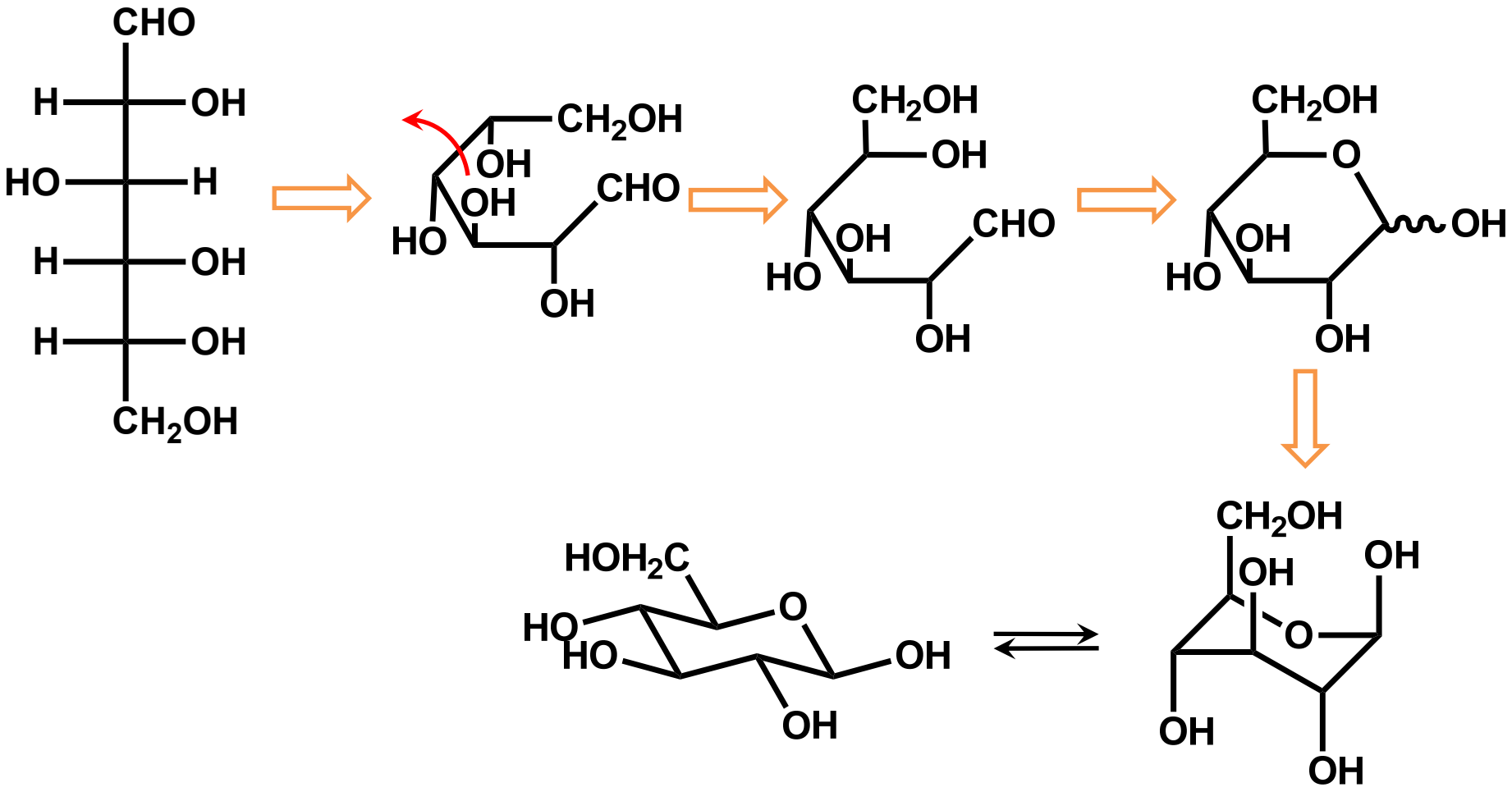


N-acetyl-D-glucosamine

- ***Other Derivatives***
  - ✓ *Glycan Esters*
    - ➢ *The –OH groups can be esterified enzymatically, e.g., phosphate esters, acyl esters, and sulfate esters*
  - ✓ *Deoxy Sugars*
    - ➢ *The –OH group can be replaced with –H, e.g., 2-deoxyribose (C-2 deoxy), fucose (C-6 deoxy), etc.*
  - ✓ *Methylated Sugars*
    - ➢ *The –OH group can be methylated (methyltransferases), e.g., –OMe group at C-1*

Monosaccharide Conformation

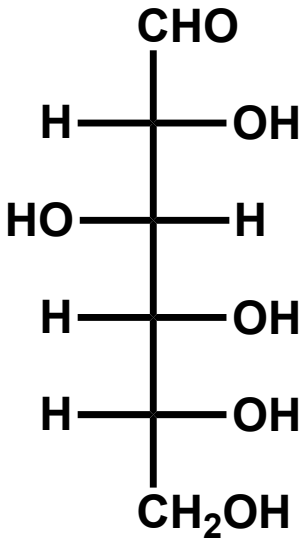
- ***Finding the Most Stable Conformation of a Monosaccharide***



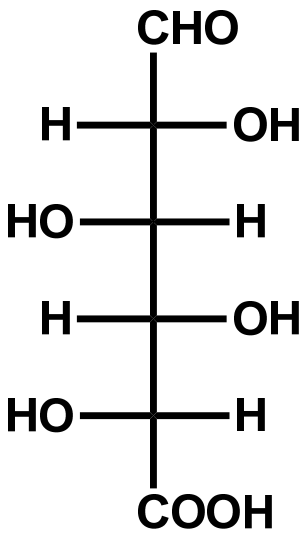
D-Glucose

Monosaccharide Conformation

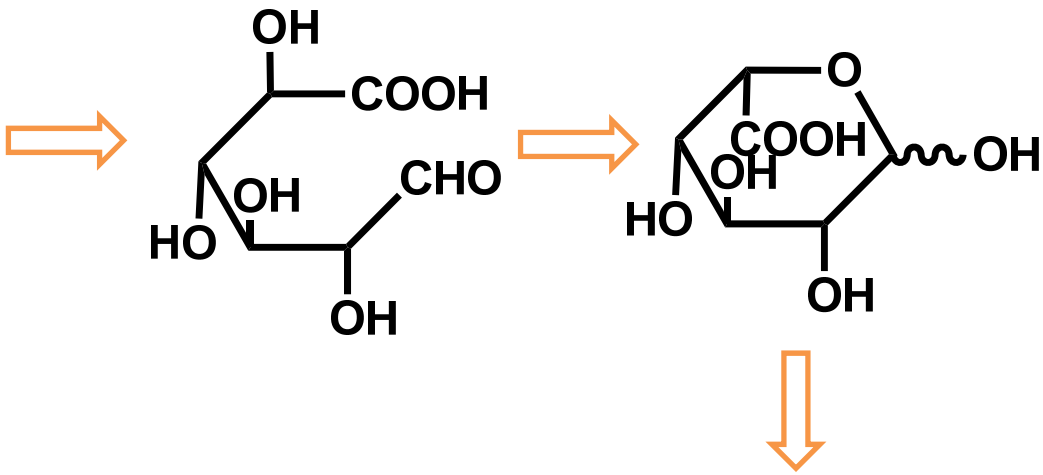
- ***Conformation of L-Iduronic Acid***

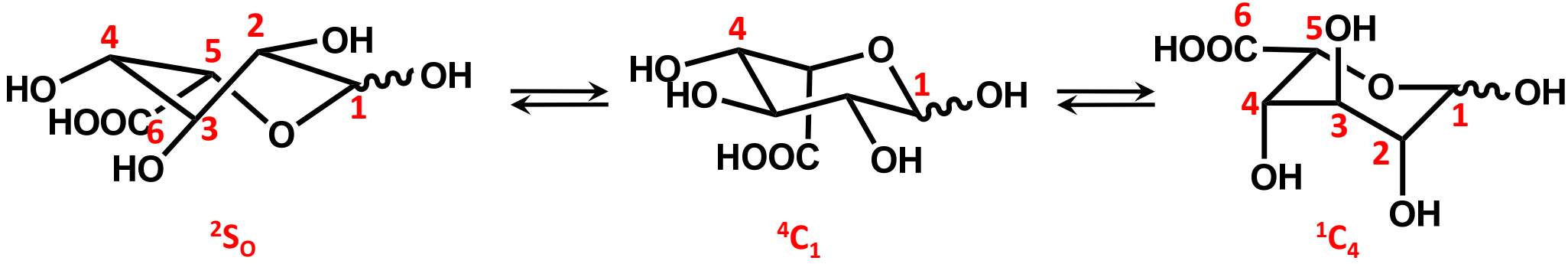


D-Glucose



L-Iduronic Acid

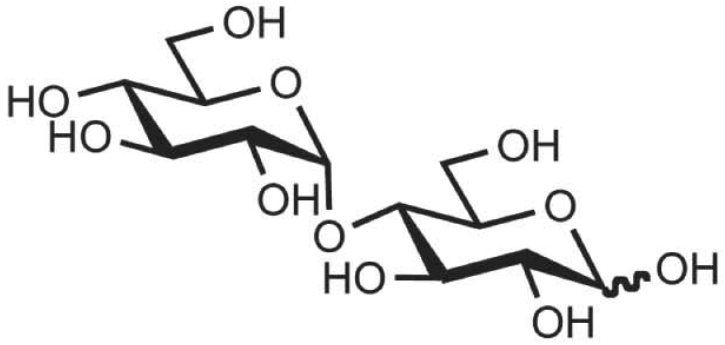




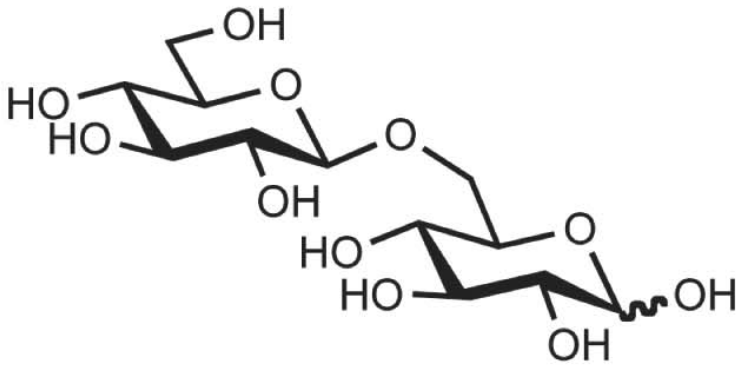
$^{2}S_{O}$ $^{4}C_{1}$ $^{1}C_{4}$

Generation of Oligosaccharides

- ***The Inter-Glycosidic Bond***
  - ✓ *The bond between two monosaccharides is the inter-glycosidic bond*
  - ✓ *Fundamental linkage for generation of oligosaccharides*
  - ✓ *Any hydroxylicated molecule (e.g., amino acids Tyr, Ser, Thr; OR alcohols cholesterol, cholic acid; OR saccharides) can form glycosidic bond*
  - ✓ *Most oligosaccharides form glycosidic bond between the anomeric carbon (C-1) and other carbons (C-2, C-3, C-4, and C-6).*
  - ✓ *These different glycosidic linkages enhance the diversity of glycans*
  - ✓ *Glycosidic bond induces directionality in the saccharide chain, i.e., non-reducing end → reducing end*
  - ✓ *Glycosidic bonds make the oligomer flexible*



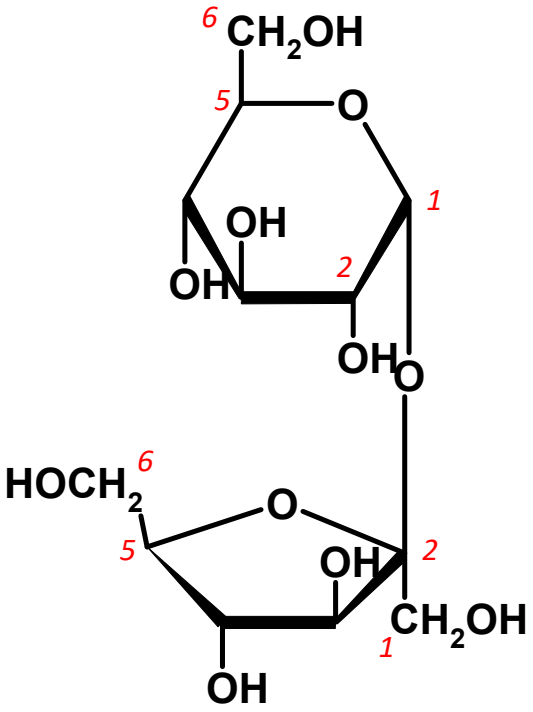
Glcα1-4Glc
(maltose)



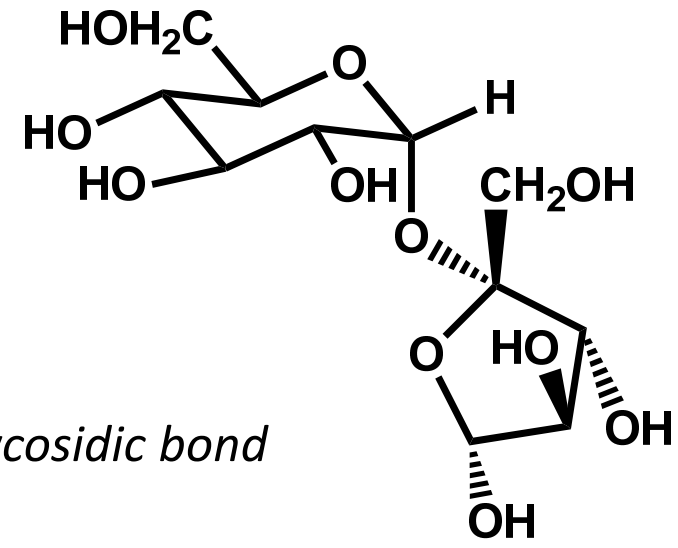
Glcβ1-6Glc
(gentiobiose)

Generation of Oligosaccharides

- ***Sucrose's Inter-Glycosidic Bond is Different***
  - ✓ *Sucrose (table sugar) is a disaccharide made from α-D-glucose and β-D-fructose, where the inter-glycosidic bond is α-1,2*



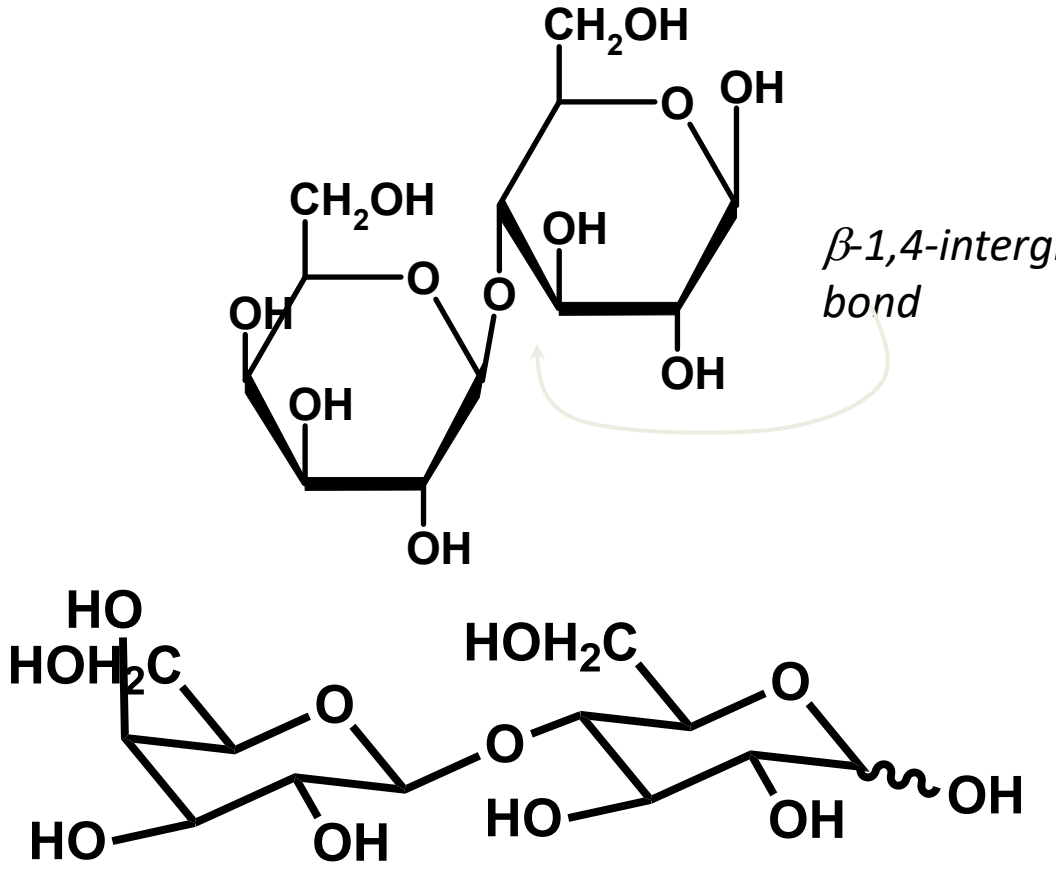
α-1,2-interglycosidic bond



  - ✓ *How do we find out which atoms are connected in polysaccharides and what is their orientation?*
  - ✓ *How do we know what conformation does a sugar adopt?*

Generation of Oligosaccharides

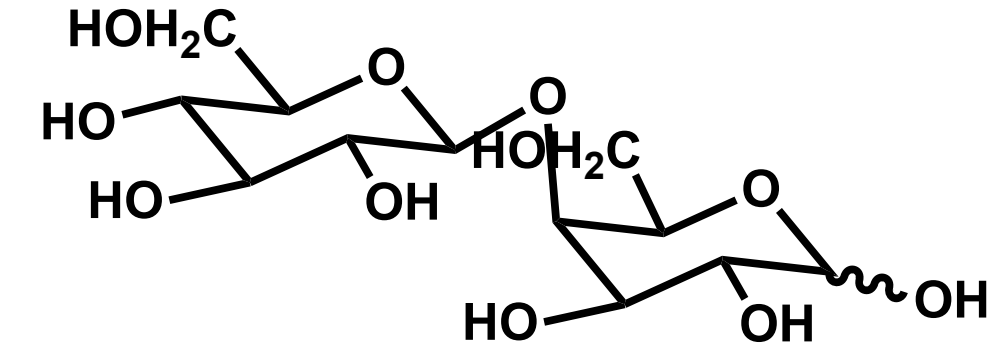
- ***Lactose***
  - ✓ *Lactose (milk sugar) is a disaccharide made from β-D-galactose and D-glucose, where the inter-glycosidic bond is β-1,4*

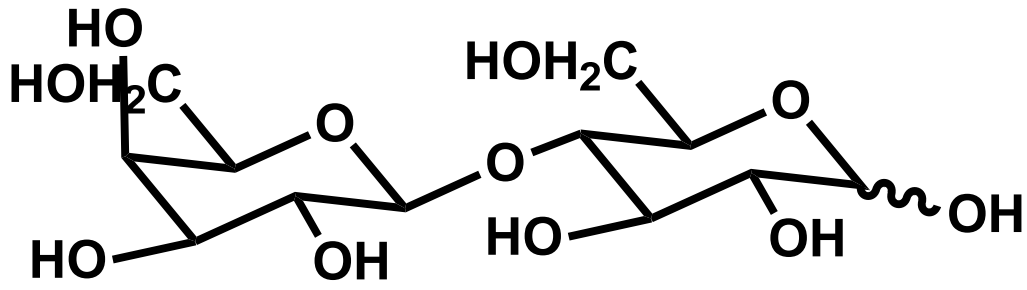


  - ✓ *Lactose intolerance? … cow's milk … 4 – 6% lactose … some infants, most blacks and many Orientals have low level of lactase, which hydrolyzes the β(1,4) bond of lactose … lactose accumulates in colon where bacterial fermentation produces large quantities of CO_2, H_2 and organic acids*

Generation of Oligosaccharides

- ***Sequence Specificity and Recognition***
  - ✓ *Is lactose (D-galactose-β (1→4)-D-glucose) is same as D-glucose-β (1→4)-D-galactose?*

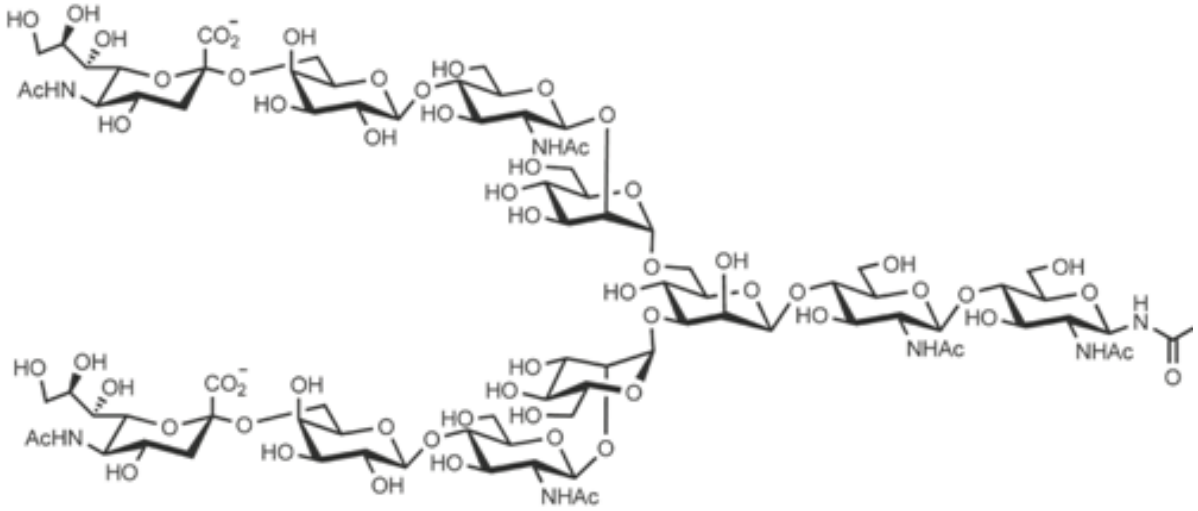




  - ✓ *Remember! D-glucose is consumed by our body; L-glucose is left untouched!*

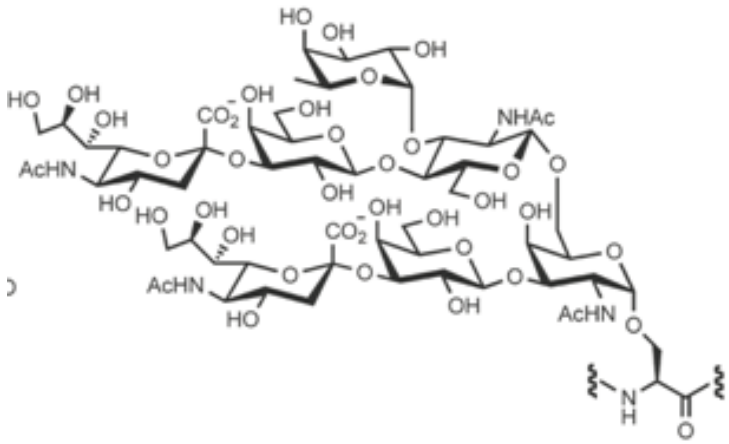
Generation of Oligosaccharides

- ***Branching is a Unique Structure Only Present in Carbohydrates***
  - ✓ *Two inter-glycosidic linkages to the same monosaccharide results in branching*
  - ✓ *Branching is not found in proteins, nucleic acids and lipids …. Only in carbohydrate*



Neu5Acα6Galβ4GlcNAcβ2Manα$_6$
Manβ4GlcNAcβ4GlcNAcβ-Asn
Neu5Acα6Galβ4GlcNAcβ2Manα$_3$

N-Glycan



Fucα$_3$
Neu5Acα3Galβ4GlcNAcβ$_6$
Neu5Acα3Galβ3GalNAcα-Ser

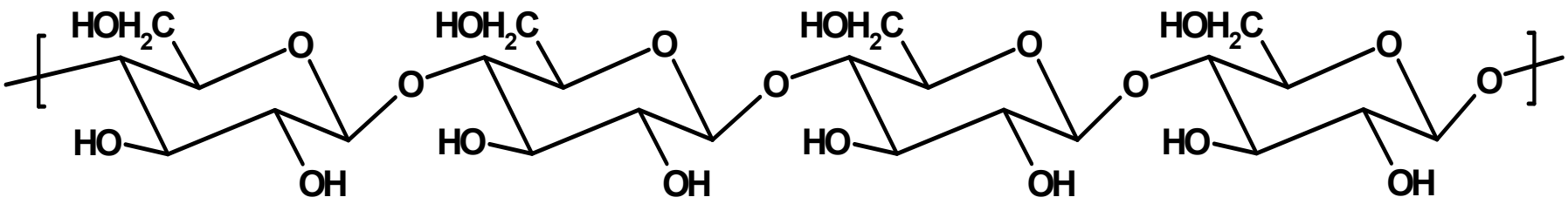
O-Glycan

Taken from 'Essentials of Glycobiology.' Edited by Varki, A.; et al.; Cold Spring Harbor Laboratory Press, New York, 2009, pg. 36.

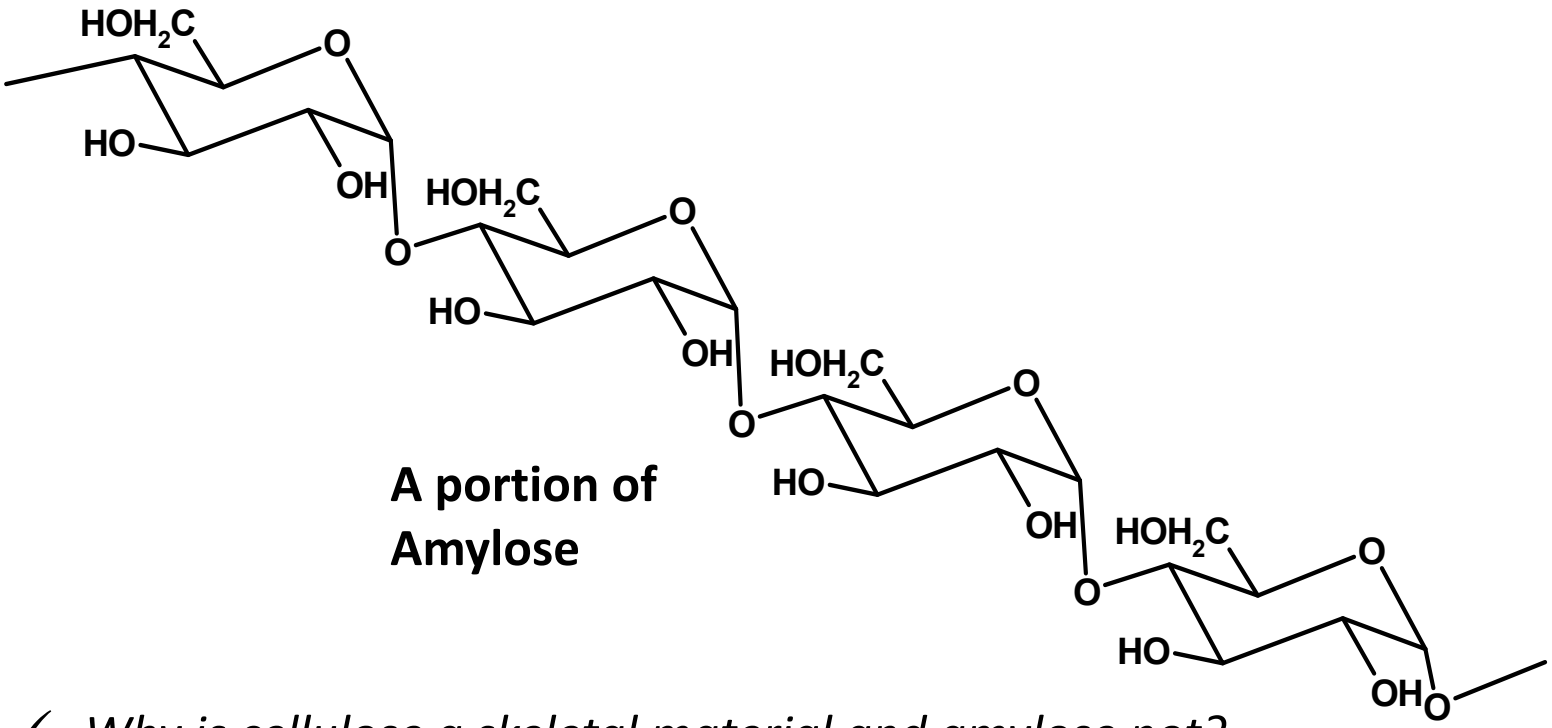
Polysaccharides

- ***Starch, cellulose & glycogen***
  - ✓ *A number of polysaccharides are known in nature. Starch, cellulose and glycogen are most common.*
  - ✓ *Starch from each plant may be different. Starch consists of two principal polysaccharides – amylose and amylopectin. Complete hydrolysis of both amylose and amylopectin yields D-glucose only. Different ways in which glucose is joined together yields different polysaccharides.*
  - ✓ *Cellulose is plant skeletal polysaccharide. Cotton is almost pure cellulose. It is a linear polysaccharide of β-1,4-linked glucose units*
  - ✓ *Glycogen is made up of α-1,4- and α-1,6-linked D-glucose units.*

Polysaccharides



cellulose

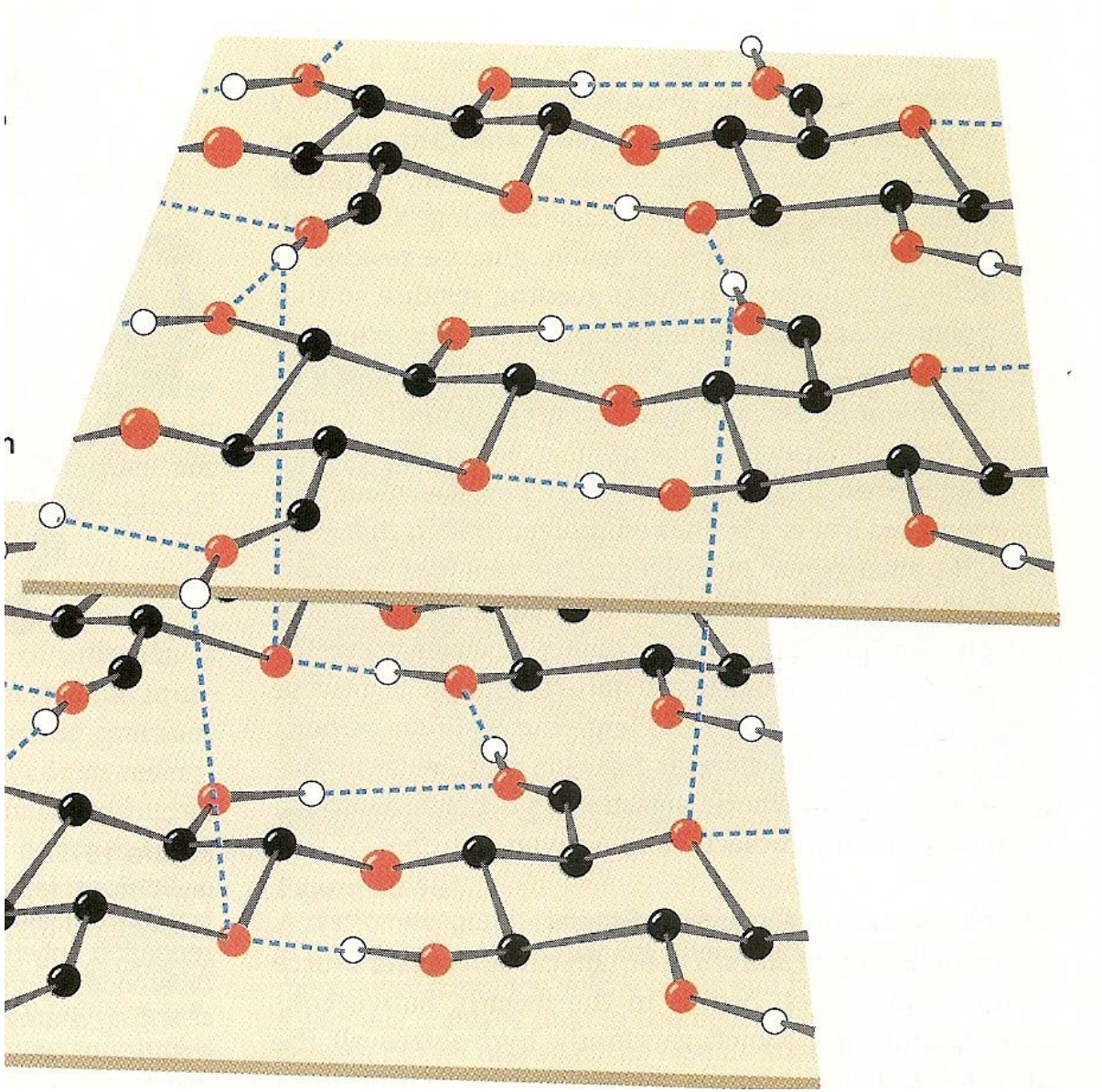


A portion of Amylose

✓ *Why is cellulose a skeletal material and amylose not?*

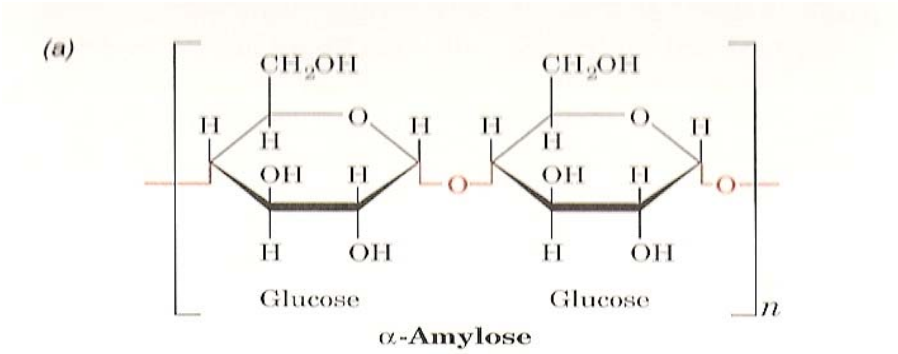
Polysaccharides

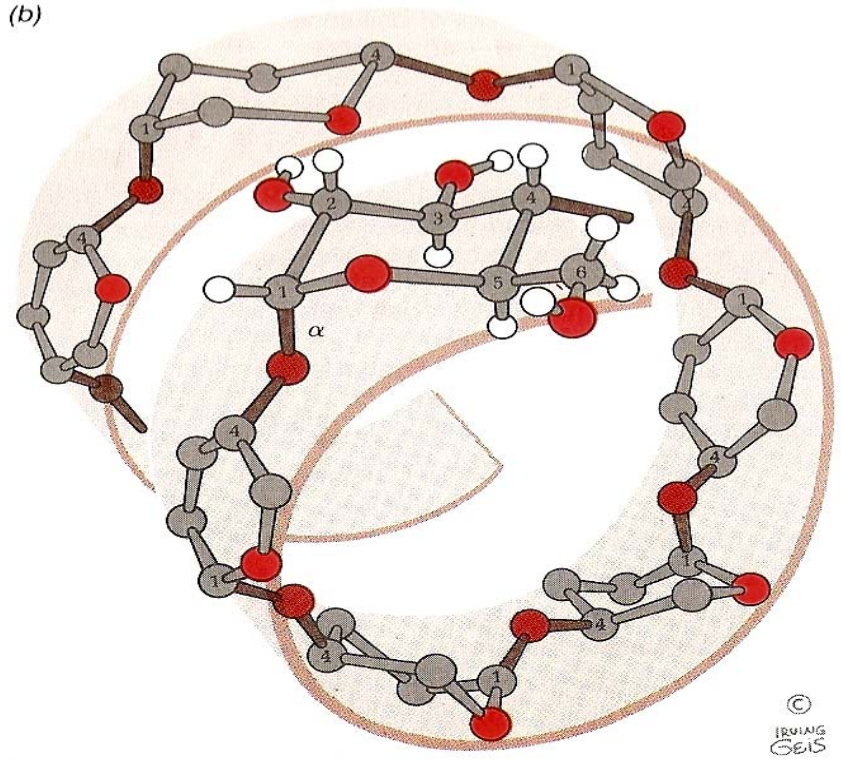
- ***Cellulose geometry affords extensive inter-chain interactions (side-by-side and one over other)***



Polysaccharides

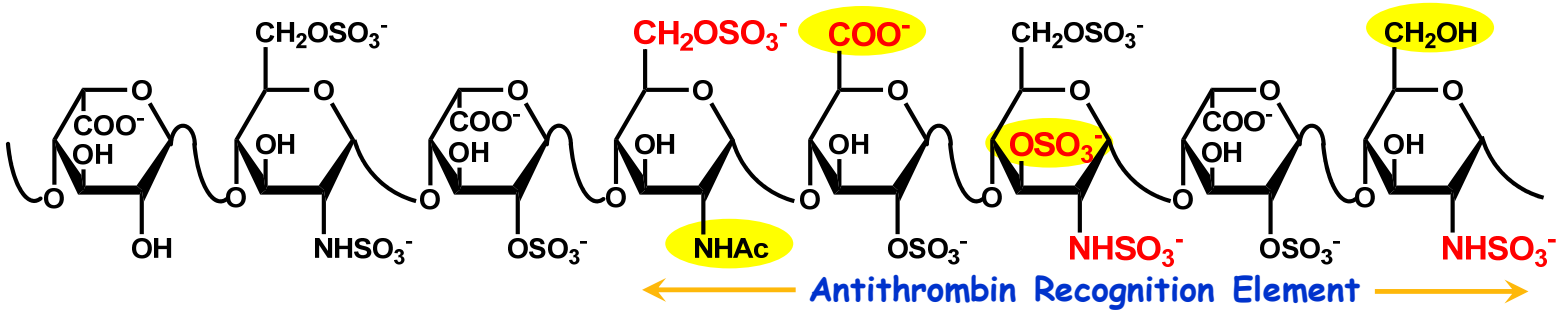
- ***Amylose (α(1→4)-linked) geometry disfavors inter-chain interactions and favors intra-chain hydrogen bonding***

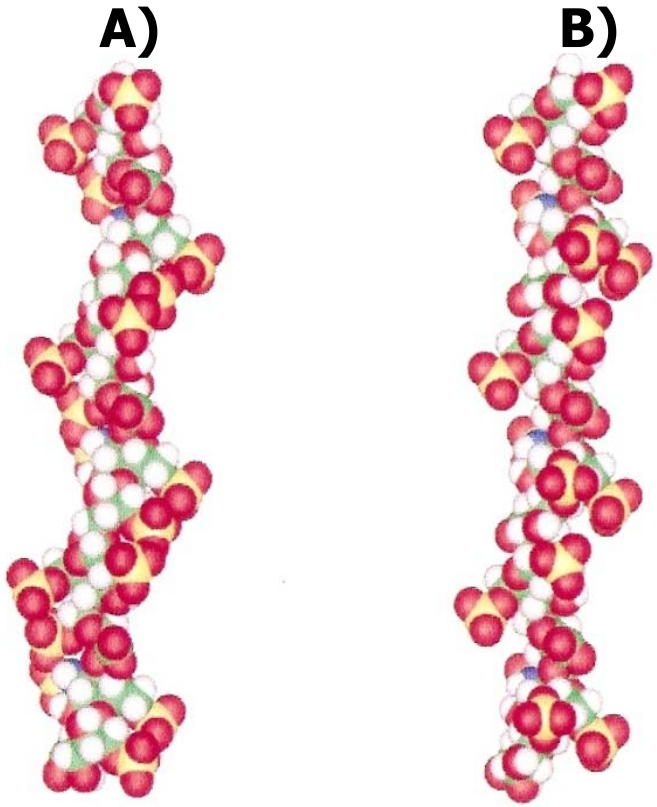




Polysaccharides

- ***Many other polysaccharides also form helical structures, e.g., heparin, dermatan sulfate, hyaluronate, …***

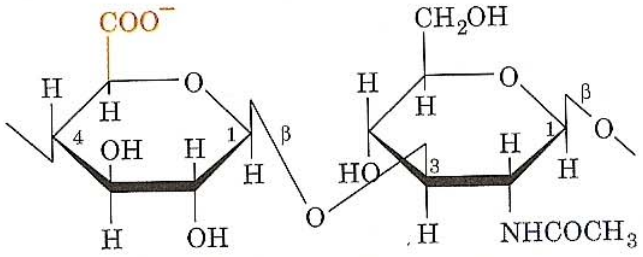




*Structure of heparin determined by NMR and molecular modeling in which each IdoAp residue is either in $^{1}C_{4}$ form (**A**) or in $^{2}S_{O}$ form (**B**) {Taken from Mulloy et al. (1993) Biochem. J. 293, 849-855}. This does not imply that all IdoAp residues are always either $^{1}C_{4}$ or $^{2}S_{O}$. See text for details.*

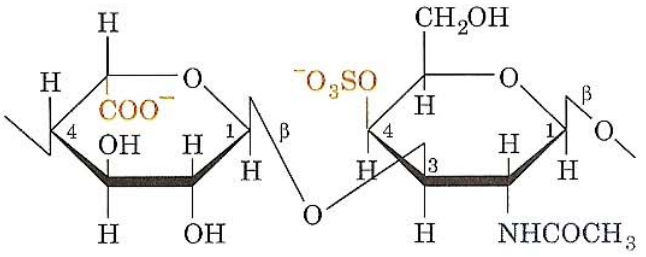
Glycosaminoglycans

- ***GAGs are heterogeneous, polydisperse, acidic polysaccharides***



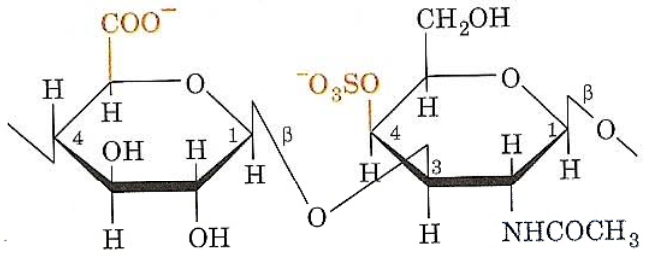
D-Glucuronurate — *N*-Acetyl-D-glucosamine

Hyaluronate



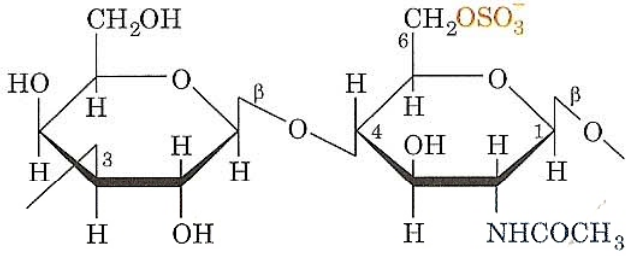
L-Iduronate — *N*-Acetyl-D-glucosamine-4-sulfate

Dermatan sulfate



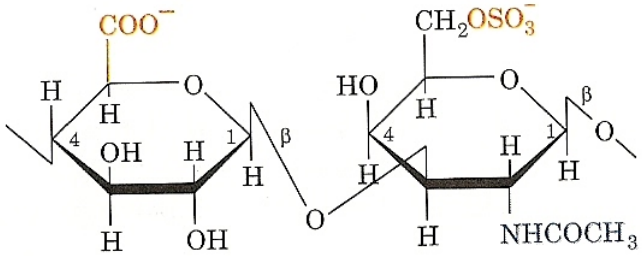
D-Glucuronurate — *N*-Acetyl-D-galactosamine-4-sulfate

Chondroitin-4-sulfate



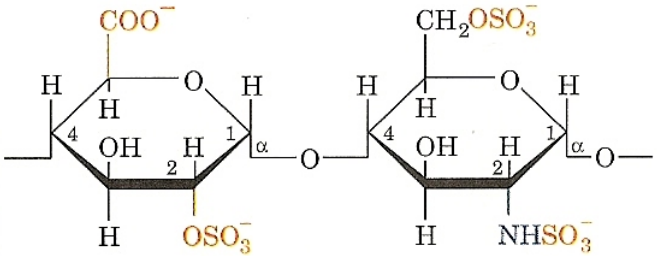
D-Galactose — *N*-Acetyl-D-glucosamine-6-sulfate

Keratan sulfate



D-Glucuronurate — *N*-Acetyl-D-galactosamine-6-sulfate

Chondroitin-6-sulfate



D-Gluconurate-2-sulfate — *N*-Sulfo-D-glucosamine-6-sulfate

Heparin

Glycosaminoglycans

- ***X-ray structure of hyaluronic Acid***

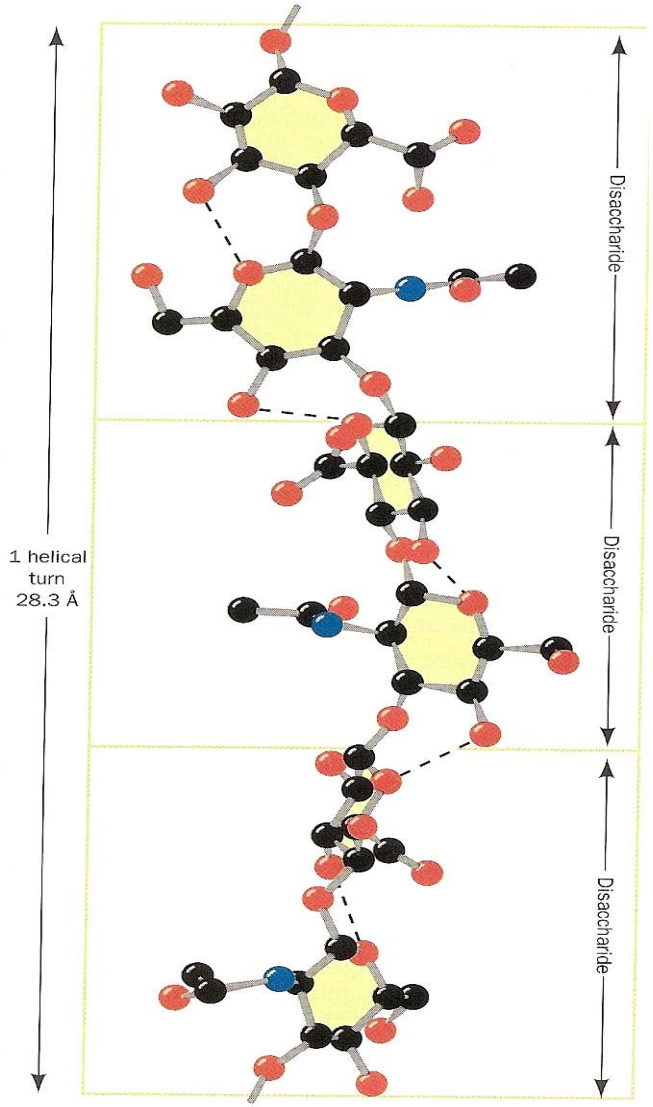


Figure 10-21
The X-ray fiber structure of Ca^{2+} hyaluronate. The hyaluronate polyanion forms an extended left-handed single-stranded helix with three disaccharide units per turn that is stabilized by intramolecular hydrogen bonds (*dashed lines*). H and Ca^{2+} atoms are omitted for clarity. [After Winter, W. T. and Arnott, S., *J. Mol. Biol.* **117,** 777 (1977).]

- ✓ *Why do GAGs prefer helical structure?*
- ✓ *What are the consequences of this arrangement?*