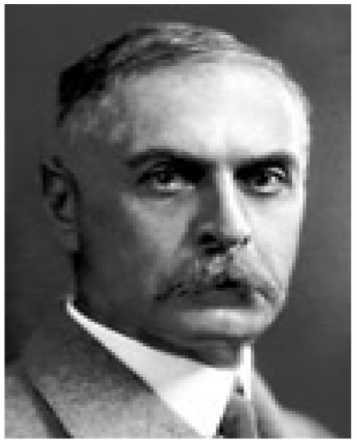
Glycobiology

The study of the biological functions, structures, recognition and biosynthesis of *glycans* (sugar chains, saccharides) in the context of the biological scaffolds to which they are attached (*e.g.* glycolipids & glycoproteins).

Nobel Laureates in Glycobiology

E. Fischer

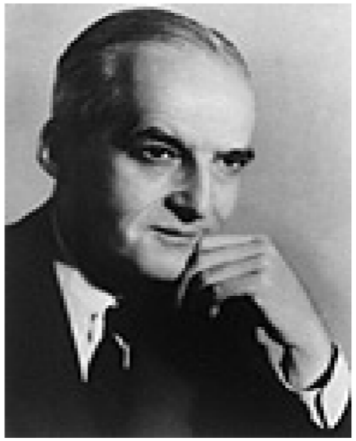


K. Landsteiner

N. Haworth

C. Cori

G. Cori

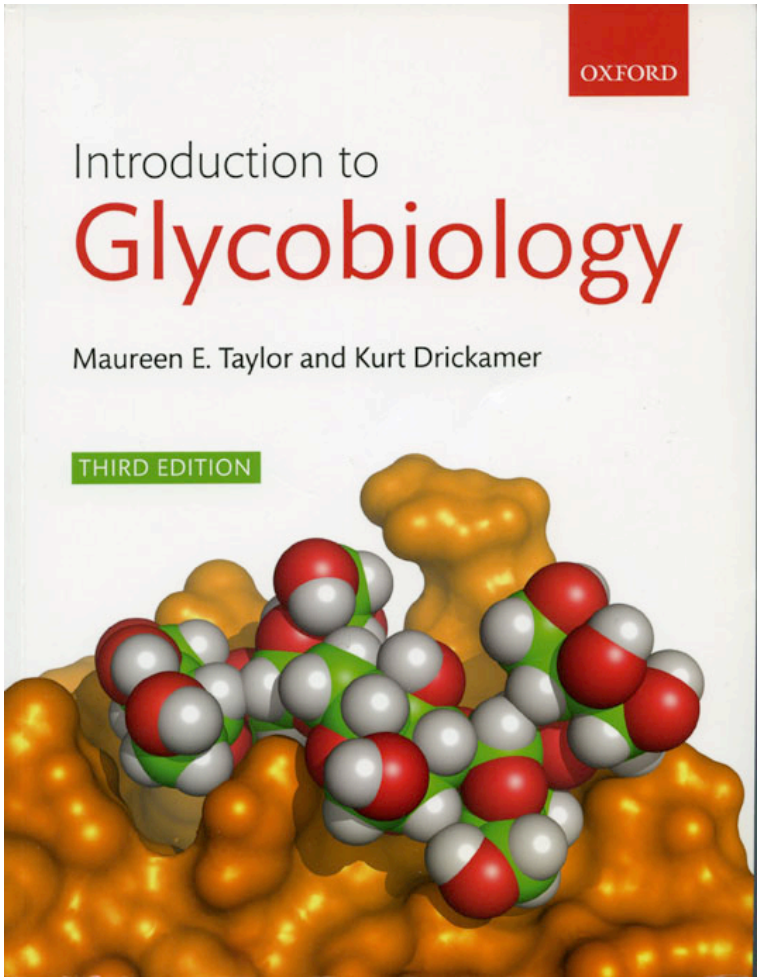


L. Leloir

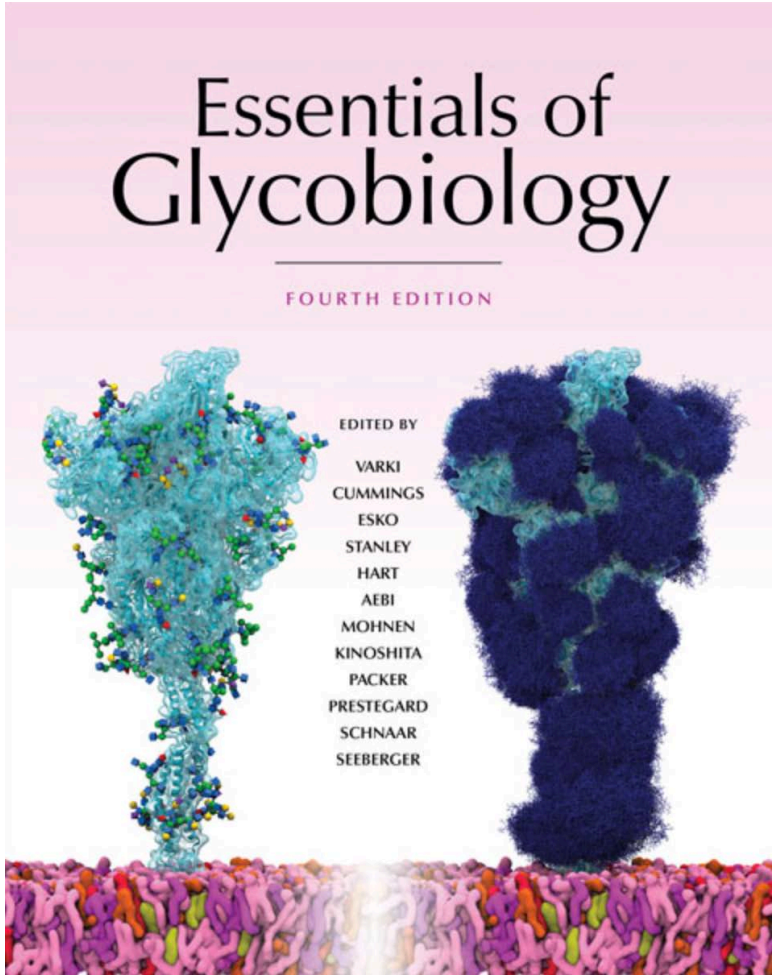
G. Palade

Also, James Rothman and Randy Schekman (2013) and Carolyn Bertozzi (2022) !!

Recommended Books on Glycobiology



Oxford University Press, 2011



CHS Press, 2022
Content freely available at:
http://www.ncbi.nlm.nih/books/NBK579918

Basic Principles of Glycobiology

Occurrence

- ✓ **All cells in nature are covered with a dense and complex array of sugar chains (glycans).**
- ✓ **The cell walls of bacteria and archea are composed of several classes of glycans and glycoconjugates.**
- ✓ **Most secreted proteins of eukaryotes also carry large amounts of covalently attached glycans.**
- ✓ **In eukaryotes, these cell-surface and secreted glycans are mostly assembled via the ER-Golgi pathway.**
- ✓ **The extracellular matrix of eukaryotes is also rich in such secreted glycans.**
- ✓ **Cytosolic and nuclear glycans are common in eukaryotes.**
- ✓ **For topological, evolutionary, and biophysical reasons, there is little similarity between cell surface/secreted and nuclear/cytosolic glycans.**

Basic Principles of Glycobiology

Chemistry and Structure

- ✓ **Glycosidic linkages can be in α- or β- linkage forms, which are biologically recognized as completely distinct**
- ✓ **Glycan chains can be linear or branched**
- ✓ **Glycans can be modified by a variety of different substituents, such as acetylation and sulfation**
- ✓ **Complete sequencing of glycans is feasible, but usually requires combinatorial or iterative methods**
- ✓ **Modern methods allow in vitro chemoenzymatic synthesis of both simple and complex glycans**

Basic Principles of Glycobiology

Biosynthesis

- ✓ **The final products of the genome are posttranslationally modified proteins, with glycosylation being the most common and versatile of these modifications.**
- ✓ **The primary units of glycans (monosaccharides) can be synthesized within a cell or salvaged from the environment.**
- ✓ **Monosaccharides must be activated into nucleotide sugars or lipid-linked sugars before they are used as donors for glycan synthesis.**
- ✓ **Whereas lipid-linked sugar donors can be flipped across membranes, nucleotide sugars must be transported into the lumen of the ER-Golgi pathway.**
- ✓ **Each linkage unit of a glycan or glycoconjugate is assembled by one or more unique glycosyltransferases.**
- ✓ **Many glycosyltransferases are members of multigene families with related functions.**
- ✓ **Most glycosyltransferases recognize the underlying glycan of their acceptor, but some are protein or lipid specific.**
- ✓ **Many biosynthetic enzymes (or glycosyltransferases, glycosidases, sulfotransferases, etc.) are expressed in a tissue-specific, temporally regulated manner.**

Basic Principles of Glycobiology

Diversity

- ✓ **Monosaccharides generate much greater combinatorial diversity than nucleotides or amino acids.**
- ✓ **Further diversity arises from covalent modifications of glycans.**
- ✓ **Glycosylation introduces a marked diversity in proteins.**
- ✓ **Only a limited subset of the potential diversity is found in a given organism or cell type.**
- ✓ **There is intrinsic diversity (microheterogeniety) within a cell type or even a single glycosylation site.**
- ✓ **The total expressed glycan repertoire (glycome) of a given cell type or organism is thus much more complex than the genome or proteome.**
- ✓ **The glycome of a given cell type or organism is also dynamic, changing in response to intrinsic and extrinsic signals.**
- ✓ **Glycome differences in cell type, space, and time generate biological diversity, and can help explain why only a limited number of genes are expressed from the typical genome.**

Basic Principles of Glycobiology

Recognition

- ✓ **Glycans are recognized by specific glycan-binding proteins that are intrinsic to an organism.**
- ✓ **Glycans are also recognized by many extrinsic glycan-binding proteins of pathogens and symbionts.**
- ✓ **Glycan-binding proteins fall in two general categories: those that can usually be grouped by shared evolutionary origins and/or similarity in structural folds (lectins) and those that emerged by convergent evolution from different ancestors (e.g, GAG-binding proteins).**
- ✓ **Lectins often show a high degree of specificity for binding to specific glycan structures, but they typically have relatively low affinities for single-site binding.**
- ✓ **Thus, biologically relevant lectin recognition usually requires multivalency of both the glycan and glycan-binding protein, to generate high avidity of binding.**

Basic Principles of Glycobiology

Genetics

- ✓ **Naturally occurring genetic defects in glycans seem to be relatively rare in intact organisms. However, this apparent rarity may be due to a failure of detection, caused by unpredictable or pleiotropic phenotypes.**
- ✓ **Genetic defects in cell-surface/secreted glycans are easily obtained in cultured cells but have somewhat limited biological consequences.**
- ✓ **The same mutations typically have major phenotypic consequences in intact multicellular organisms.**
- ✓ **Thus, many of the major roles of glycans likely involve cell–cell or extracellular interactions.**
- ✓ **Nuclear/cytosolic glycans may have more cell-intrinsic roles, e.g., in signaling.**
- ✓ **Complete elimination of major glycan classes generally causes early developmental lethality.**
- ✓ **Organisms bearing tissue-specific alteration of glycans often survive, but they exhibit both cell-autonomous and distal biological effects.**

Basic Principles of Glycobiology

Biological Roles

- ✓ **Biological roles for glycans span the spectrum from nonessential activities to those that are crucial for the development, function, and survival of an organism.**
- ✓ **Many theories regarding the biological roles of glycans appear to be correct, but exceptions occur.**
- ✓ **Glycans can have different roles in different tissues or at different times in development.**
- ✓ **Terminal sequences, unusual glycans, and modifications are more likely to mediate specific biological roles.**
- ✓ **However, terminal sequences, unusual glycans, or modifications may also reflect prior evolutionary interactions with microorganisms and other noxious agents.**
- ✓ **Thus, a priori prediction of the functions of a specific glycan or its relative importance to the organism is difficult.**

Basic Principles of Glycobiology

Evolution

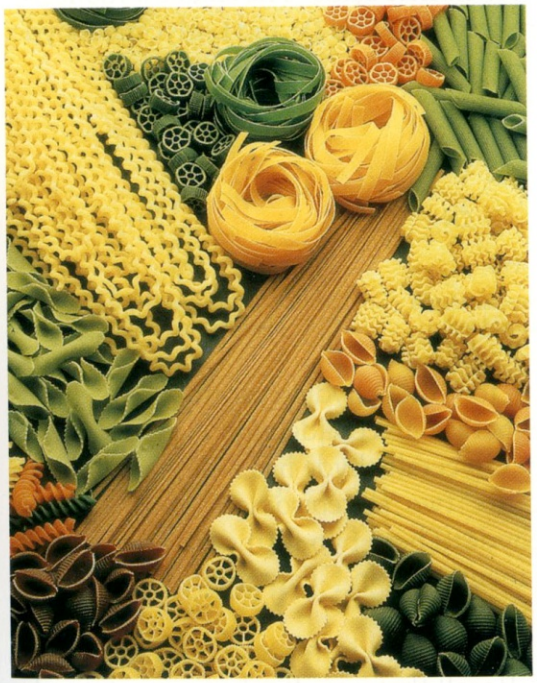
- ✓ **Relatively little is known about glycan evolution.**
- ✓ **Interspecies and intraspecies variations in glycan structure are relatively common, suggesting rapid evolution.**
- ✓ **The dominant mechanism for such evolution is likely the ongoing selection pressure by pathogens that recognize glycans.**
- ✓ **However, glycan evolution must also preserve and/or elaborate critical intrinsic functions.**
- ✓ **Interplay between pathogen selection pressure and preservation of intrinsic roles could result in the formation of "junk" glycans.**
- ✓ **Such "junk" glycans could be the substrate from which new intrinsic functions arise during evolution.**

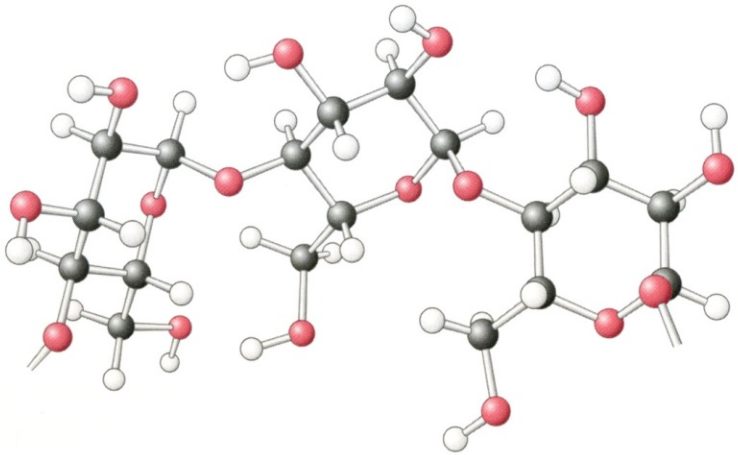
Glycobiology is …

NOT Carbohydrates as Food

Carbohydrates

Carbohydrates in food are important sources of energy. Starch found in plant-derived food such as pasta ....



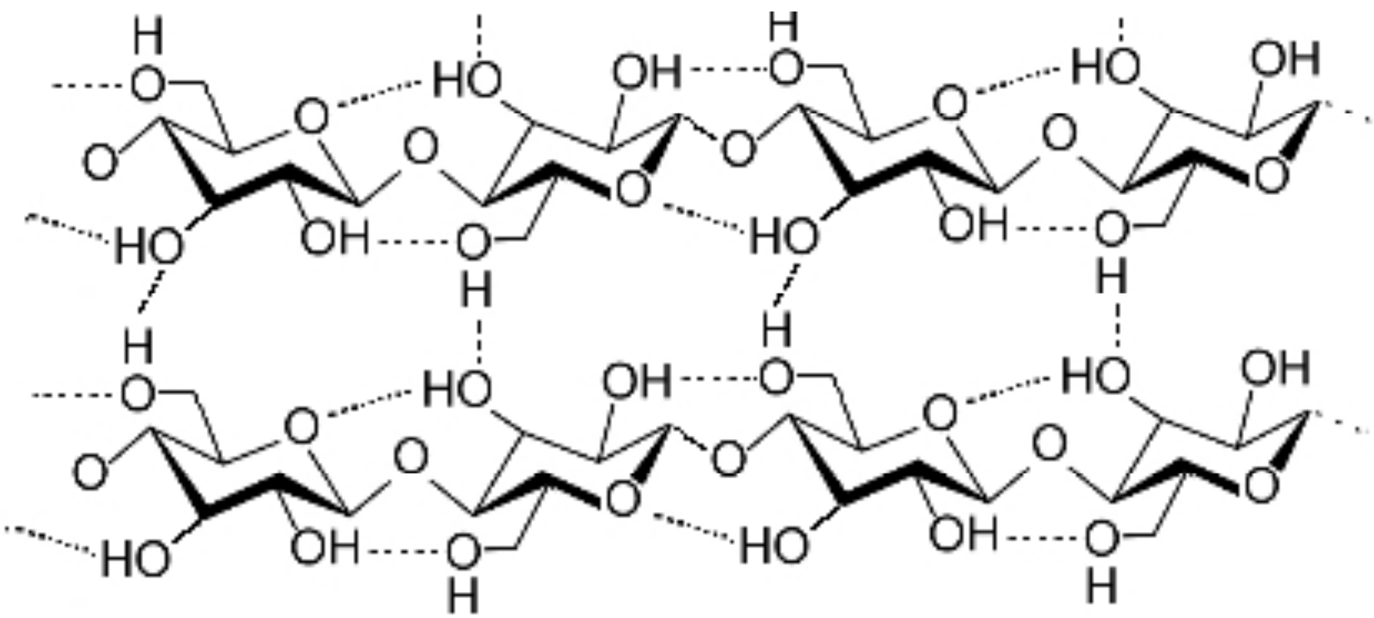


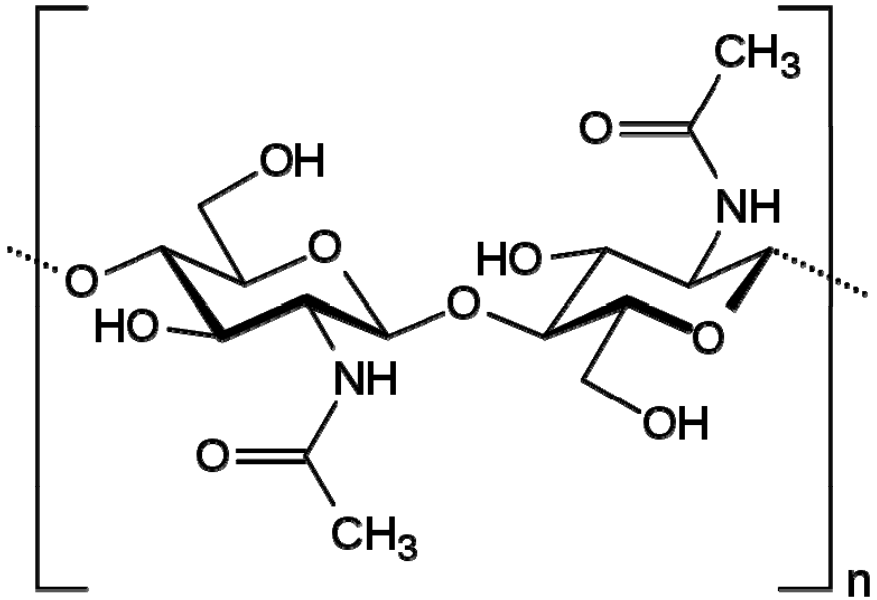
Carbohydrates in food are important sources of energy. Starch, found in plant-derived food such as pasta, consists of chains of linked glucose molecules. These chains are broken down into individual glucose molecules for eventual use in the generation of ATP and as building blocks for other molecules. [(Left) Superstock.]

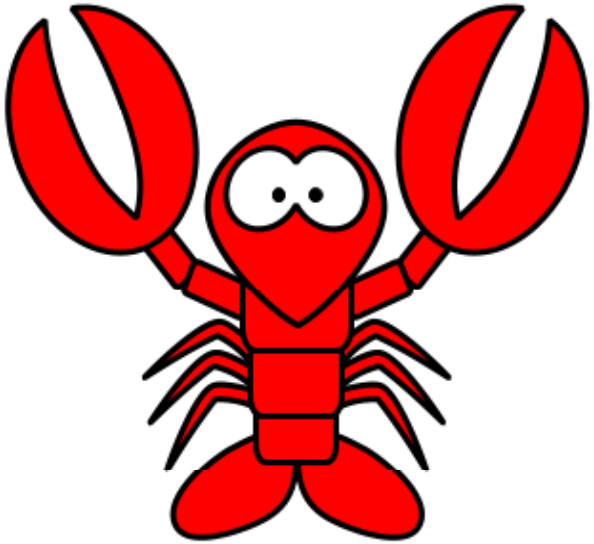
Berg, Tymoczko & Stryer (2006) *Biochemistry, Sixth Edition*, W.H. Freeman, New York

Cellulose

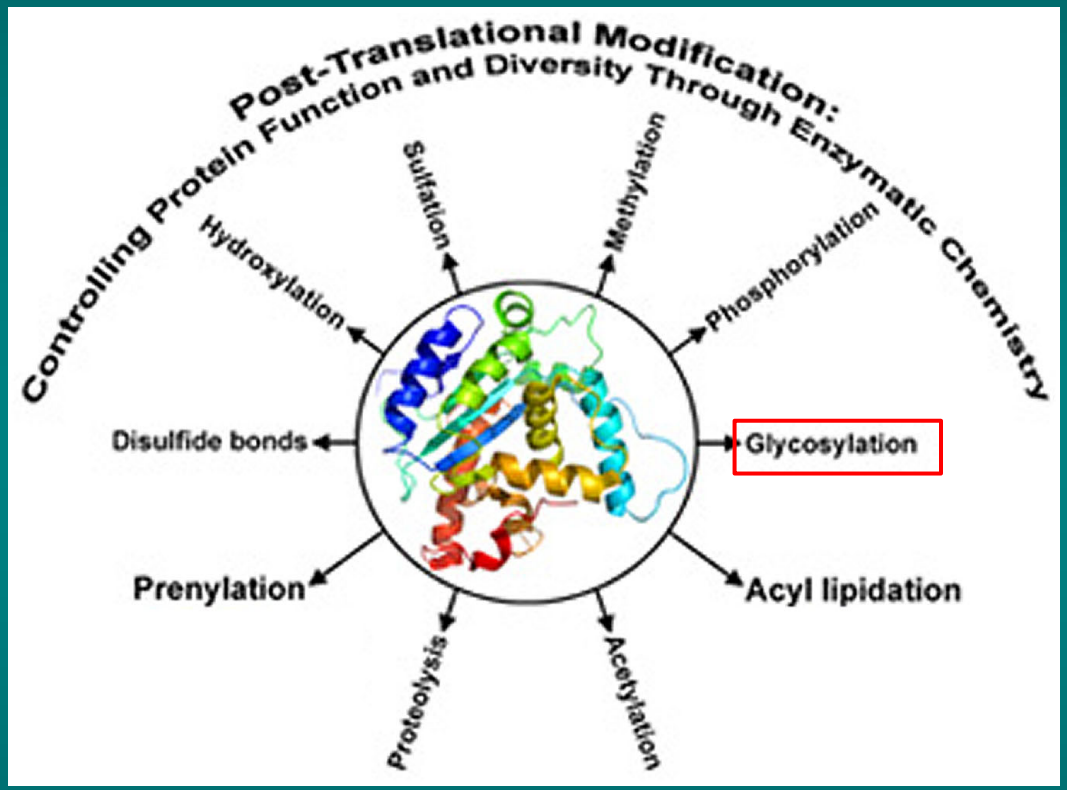






Chitin

Glycobiology is NOT "a" Post-Translational Modification…. it is Thousands of Modifications



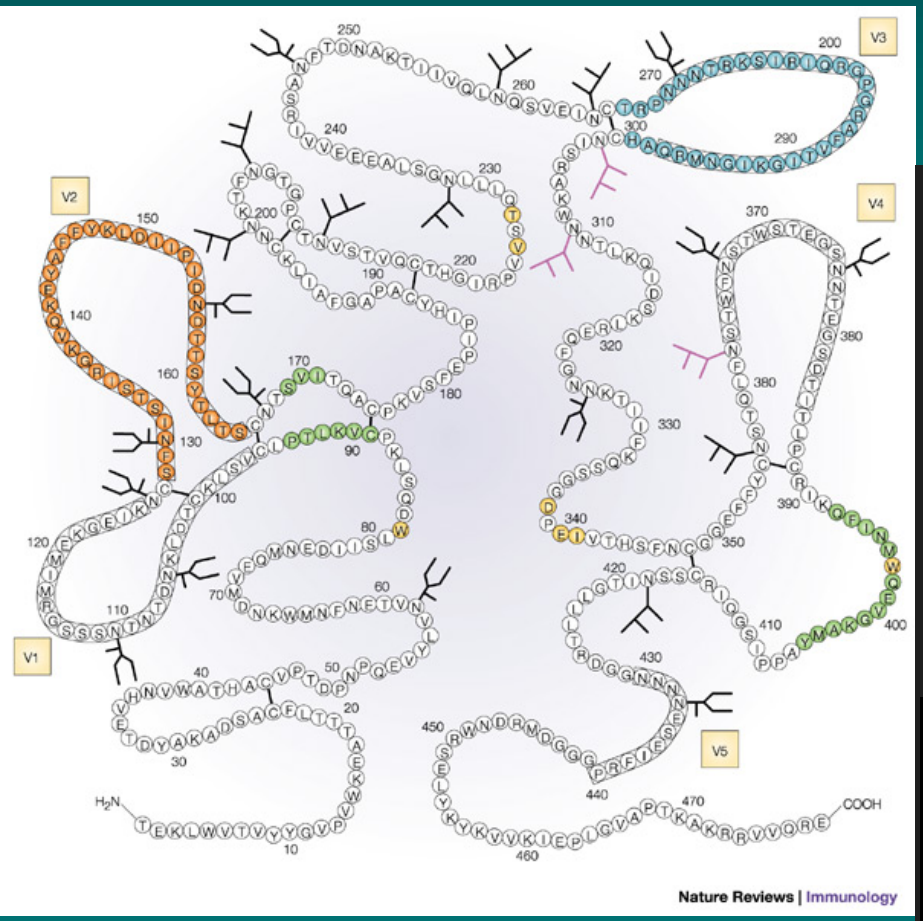
Modified from http://www-che.syr.edu/ faculty/hougland.html

- Combinatorial linkage position, orientation and branching provide the potential for millions* of different glycan structures – functional diversity
- Glycans may be larger and are more diverse than their (protein) carriers
- A glycan's function can supersede that of its (protein) carrier
- 1-2% of the human genome is devoted to glycosylation

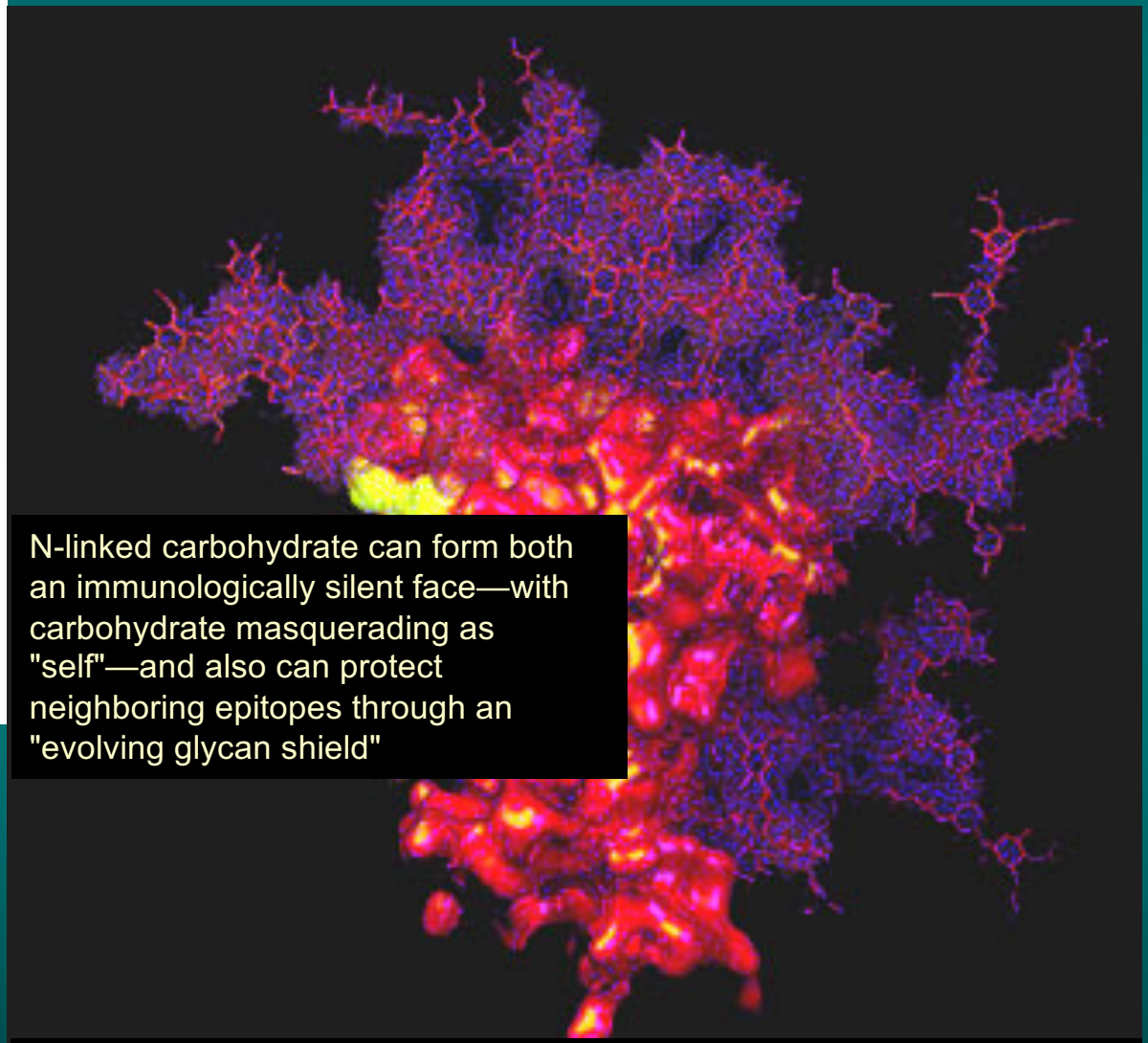
*Glycoproteins and glycolipids may contain ~3000 glycan determinants with an additional ~4000 theoretical pentasaccharide sequences in glycosaminoglycans

Cummings RD (2009) *Molecular BioSystems* 5, 1087

Glycobiology is … NOT a "Decoration"



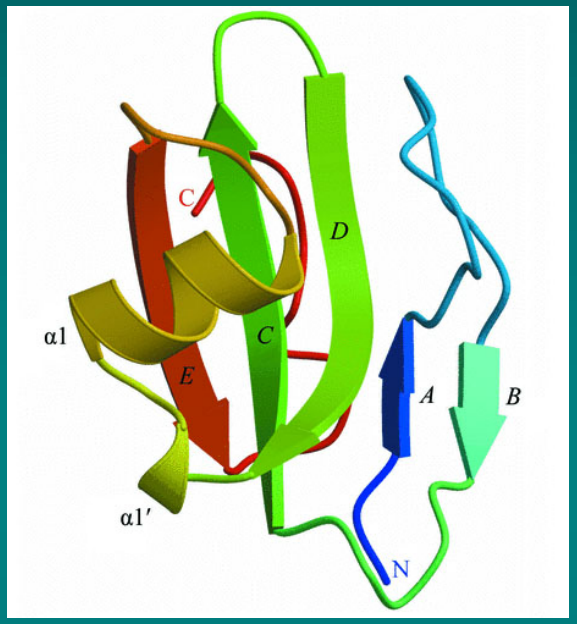
"Structure" of HIV gp120
Zolla-Pazner (2004)
Nature Reviews Immunology 4, 199



N-linked carbohydrate can form both an immunologically silent face—with carbohydrate masquerading as "self"—and also can protect neighboring epitopes through an "evolving glycan shield"

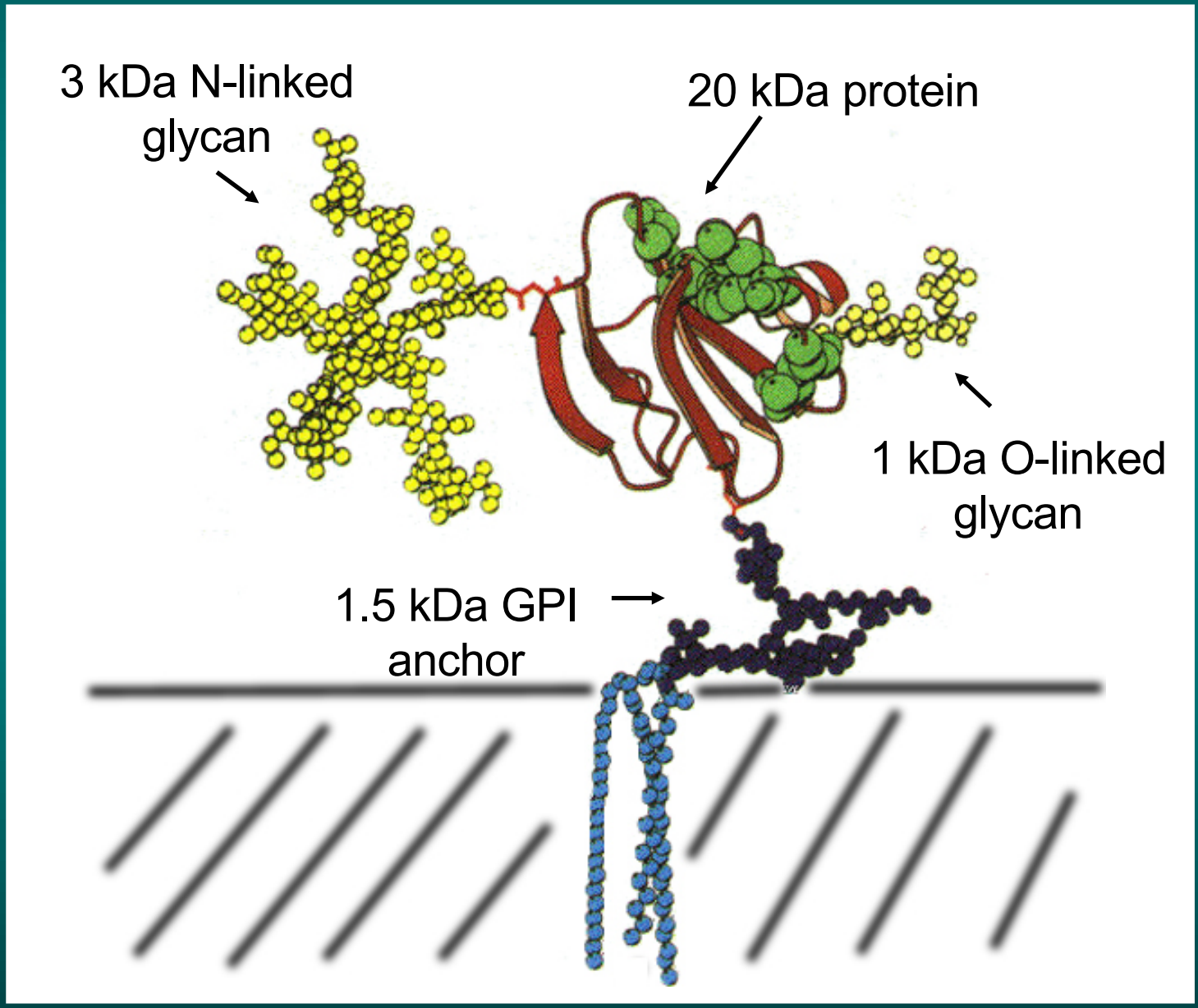
http://www.niaid.nih.gov/labsandresources/labs/aboutlabs/vrc/structuralbiologylaboratory/Pages/kwong.aspx

Proteins typically fold inward, whereas glycans spread out in space



"naked" CD59, Huang et al (2007) *Acta Crystallographica* 63, 714

CD59, a complement defense glycoprotein



Rudd et al (1997) *J Biol Chem* 272, 7229

Basic Definitions

- **Monosaccharide:** A carbohydrate that cannot be hydrolyzed into a simpler carbohydrate. The building block of oligosaccharides and polysaccharides.
- **Oligosaccharide:** Linear or branched chain of monosaccharides attached to one another via glycosidic linkages. The number of monosaccharide units can vary.
- **Polysaccharide:** Glycan composed of repeating monosaccharides, generally greater than ten monosaccharide units in length.
- **Carbohydrate, glycan, saccharide, sugar:** Generic terms used interchangeably. Includes monosaccharides, oligosaccharides, polysaccharides, and derivatives of these compounds. Carbohydrates consist of "hydrated carbon", $[CH_2O]_n$
- Preferred generic term is "Glycan"

Glycoconjugates

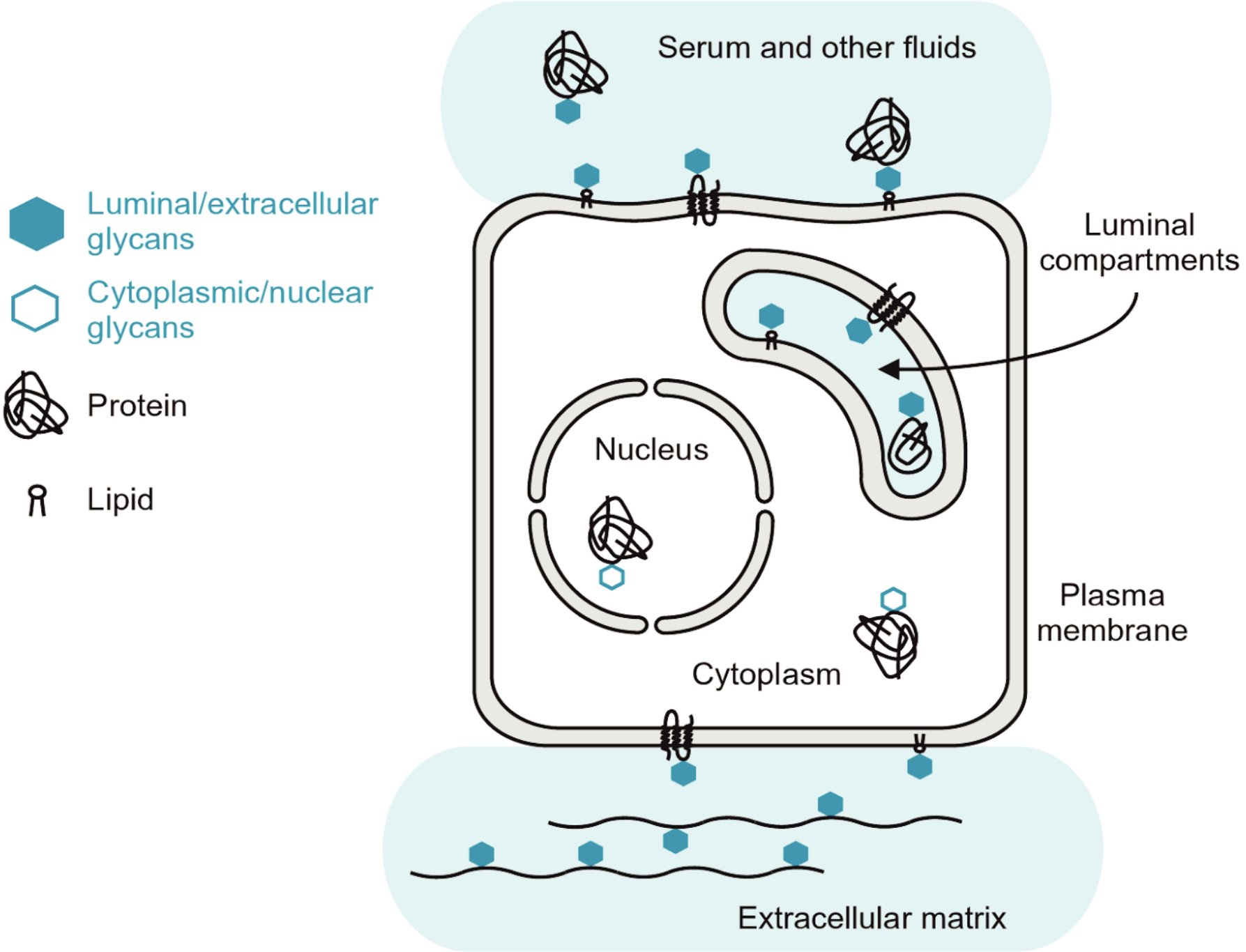
Glycoconjugate: **A compound in which one or more glycans (the glycone) are covalently linked to a non-carbohydrate moiety (the aglycone)**

Glycoproteins: **Protein with one or more covalently bound glycans**

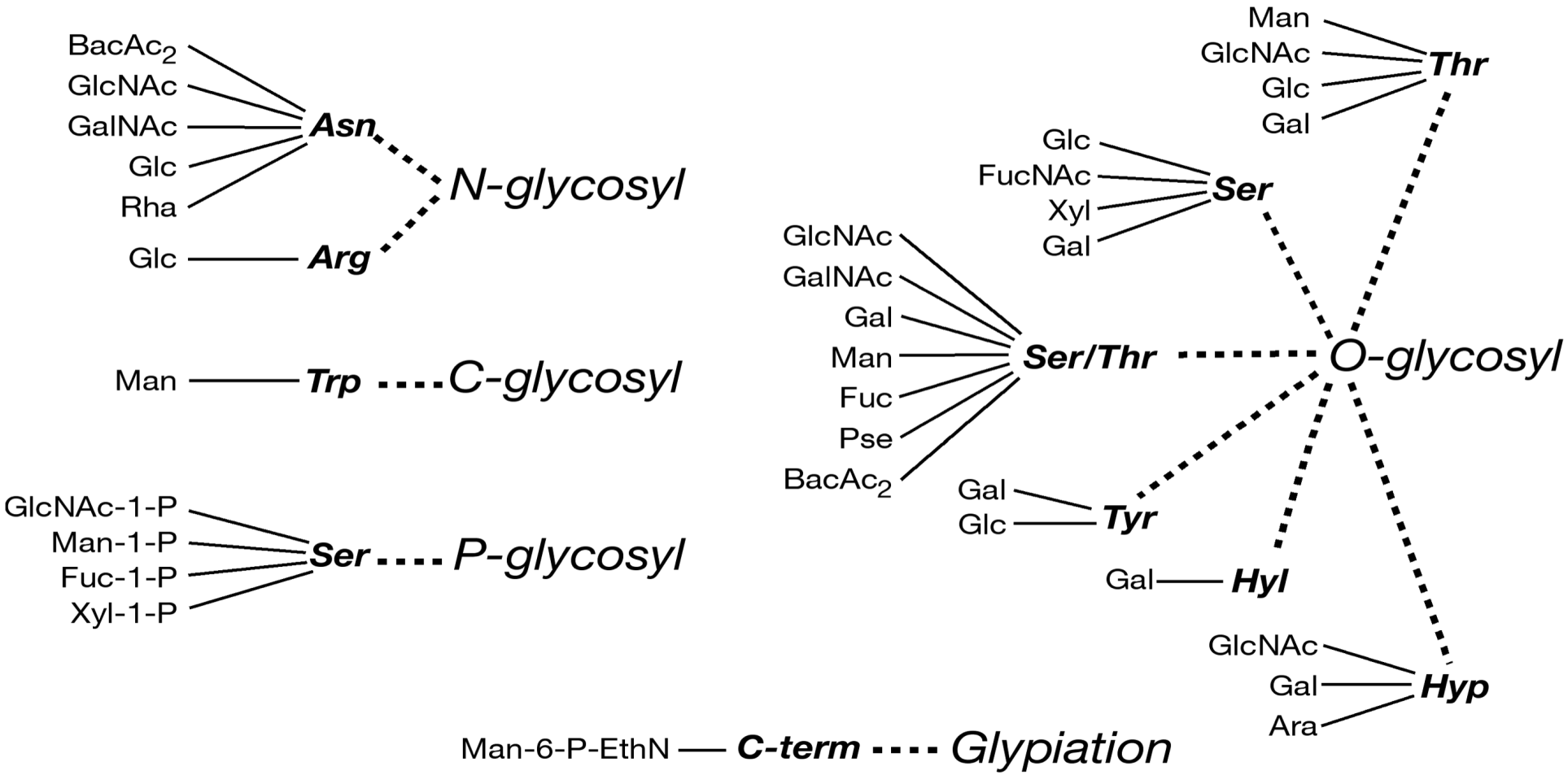
Glycolipids: **A molecule containing a saccharide linked to a lipid**

Proteoglycans: **Any glycoprotein with one or more covalently attached glycosaminoglycan chains**

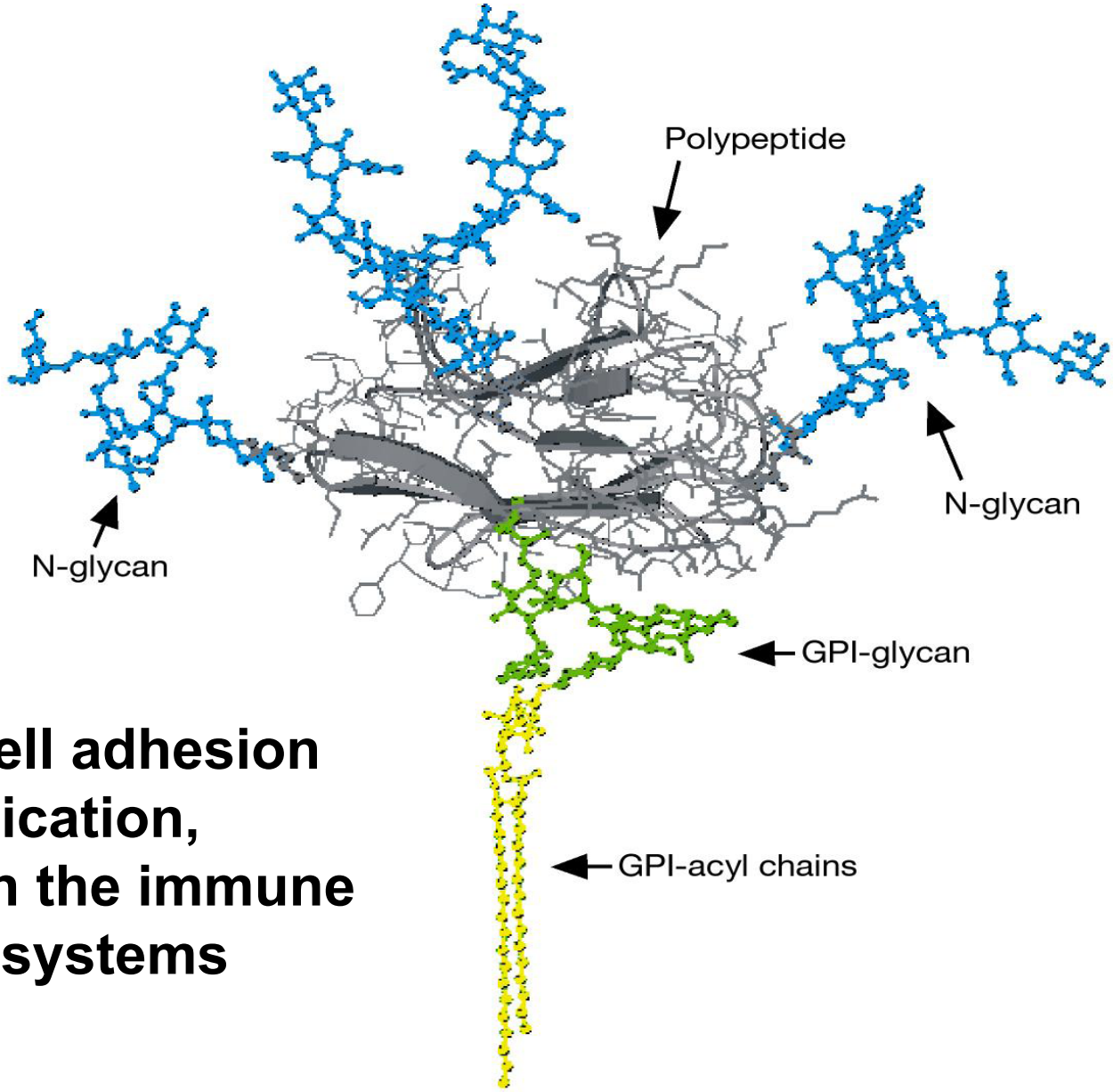
Intra- & Extracellular Localization of Glycoconjugates



Types of Glycan-Protein Linkages in Nature

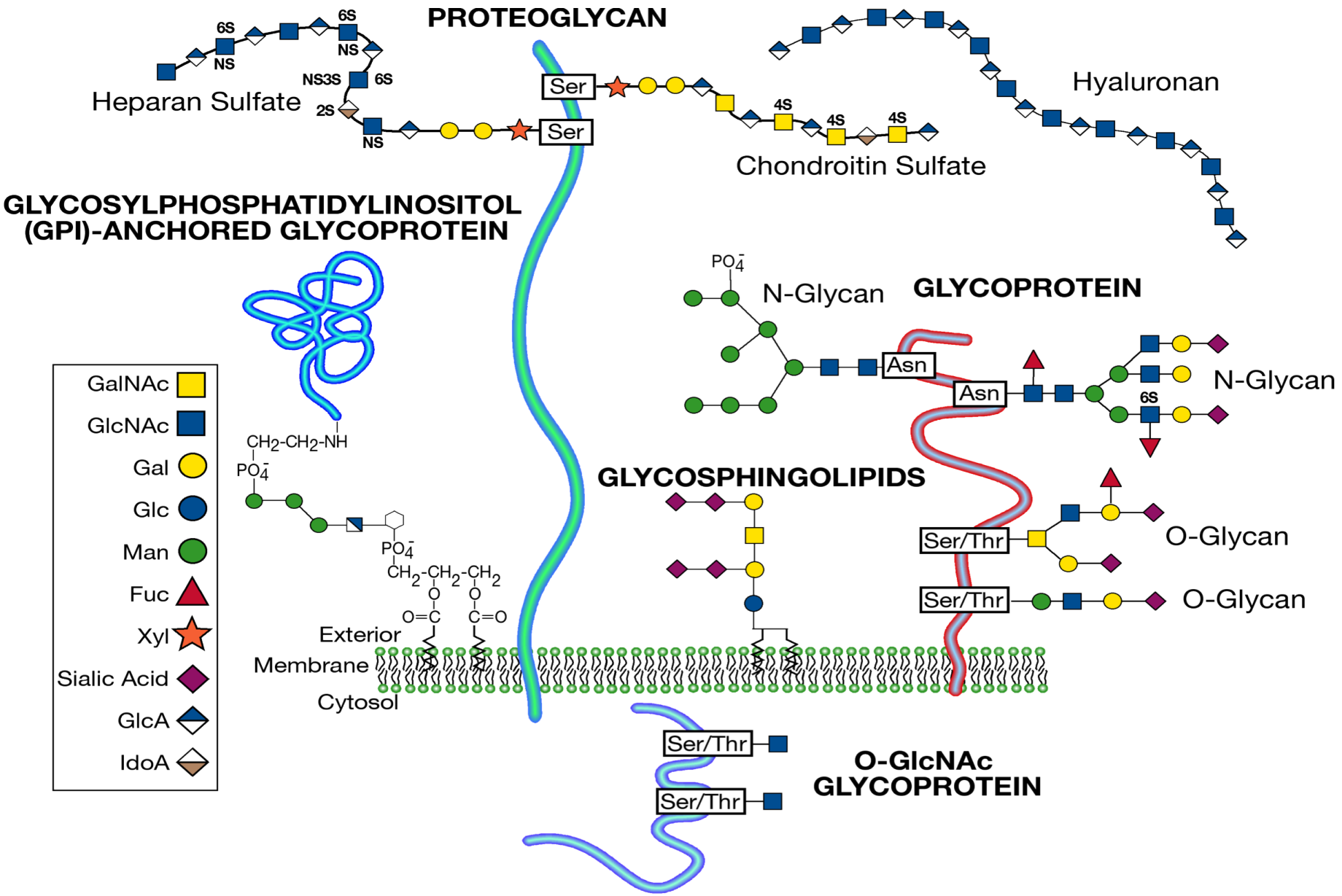


Thy-1 Glycoprotein – A Cell Surface Antigen

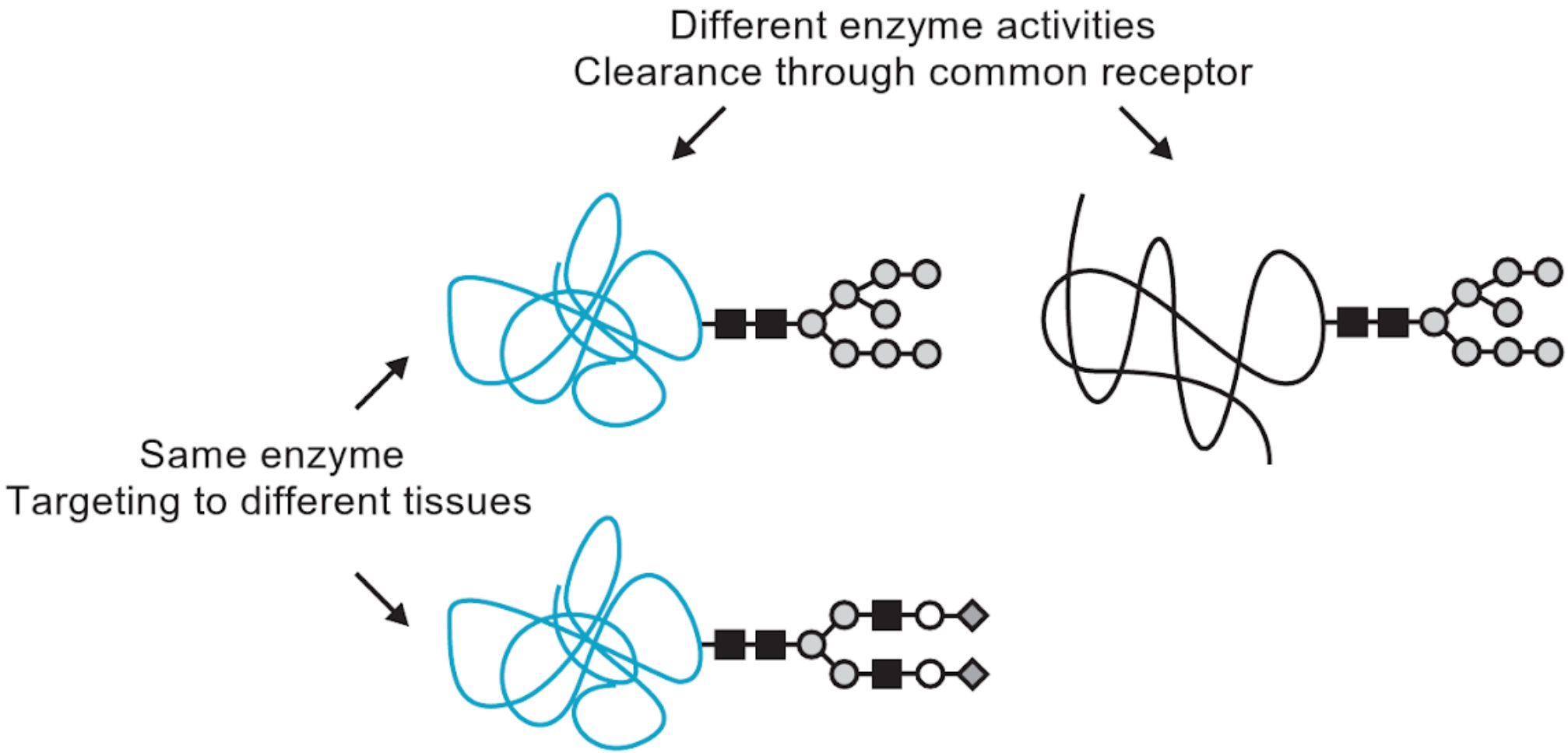


Involved in cell adhesion and communication, particularly in the immune and nervous systems

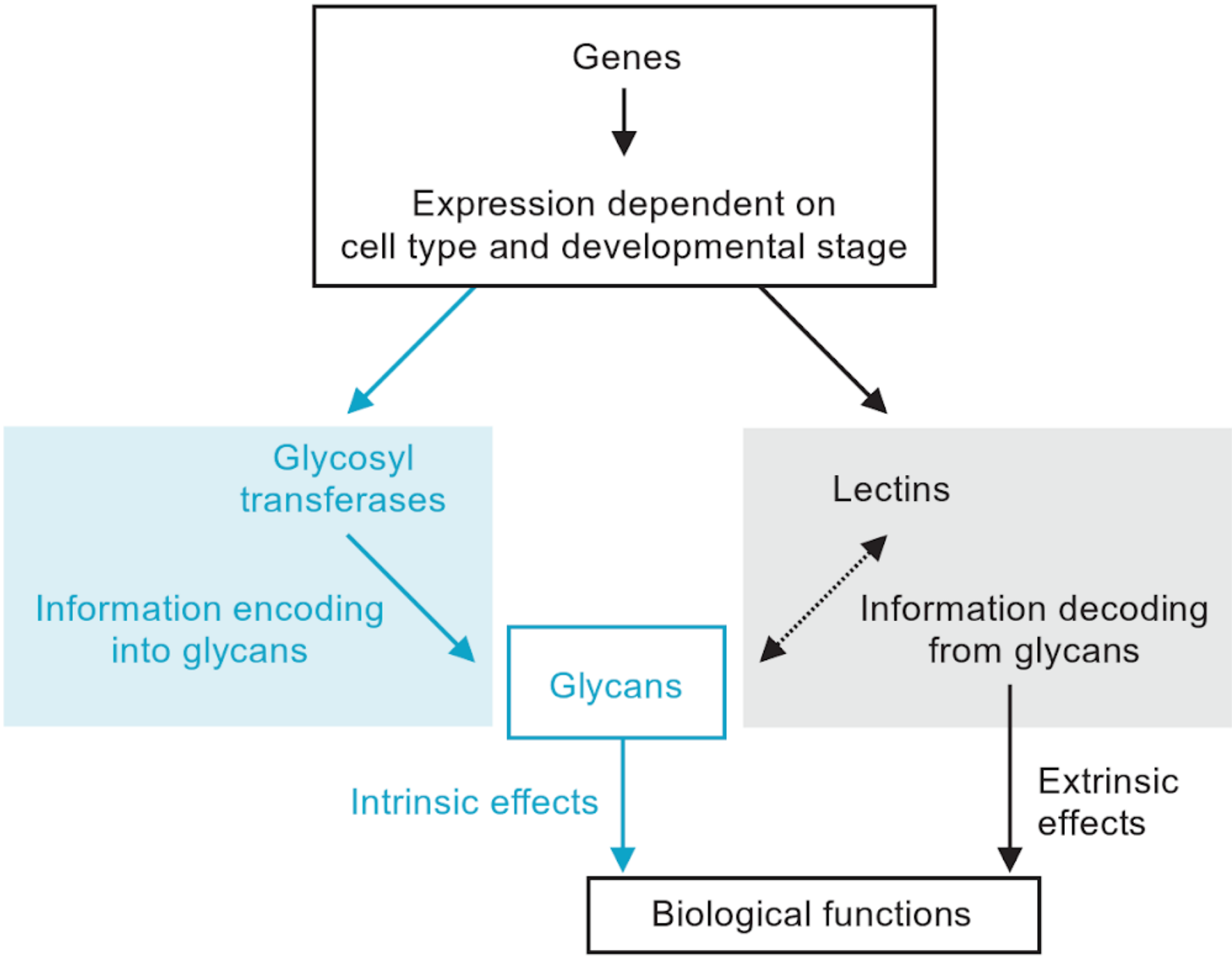
Major Glycan Classes in Vertebrate Cells



Independent Functions of Protein & Glycan



Encoding & Decoding Information in Glycans



Summary of Glycan Functions

Providing structural components
- Cell walls
- Extracellular matrix

Modifying protein properties
- Solubility
- Stability

Intrinsic functions performed by glycans

Directing trafficking of glycoconjugates
- Intracellular
- Extracellular

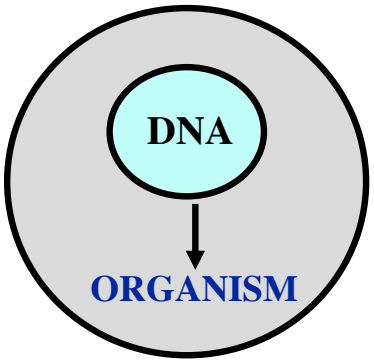
Mediating and modulating cell adhesion
- Cell–cell interactions
- Cell–matrix interactions

Mediating and modulating signalling
- Intracellular
- Extracellular

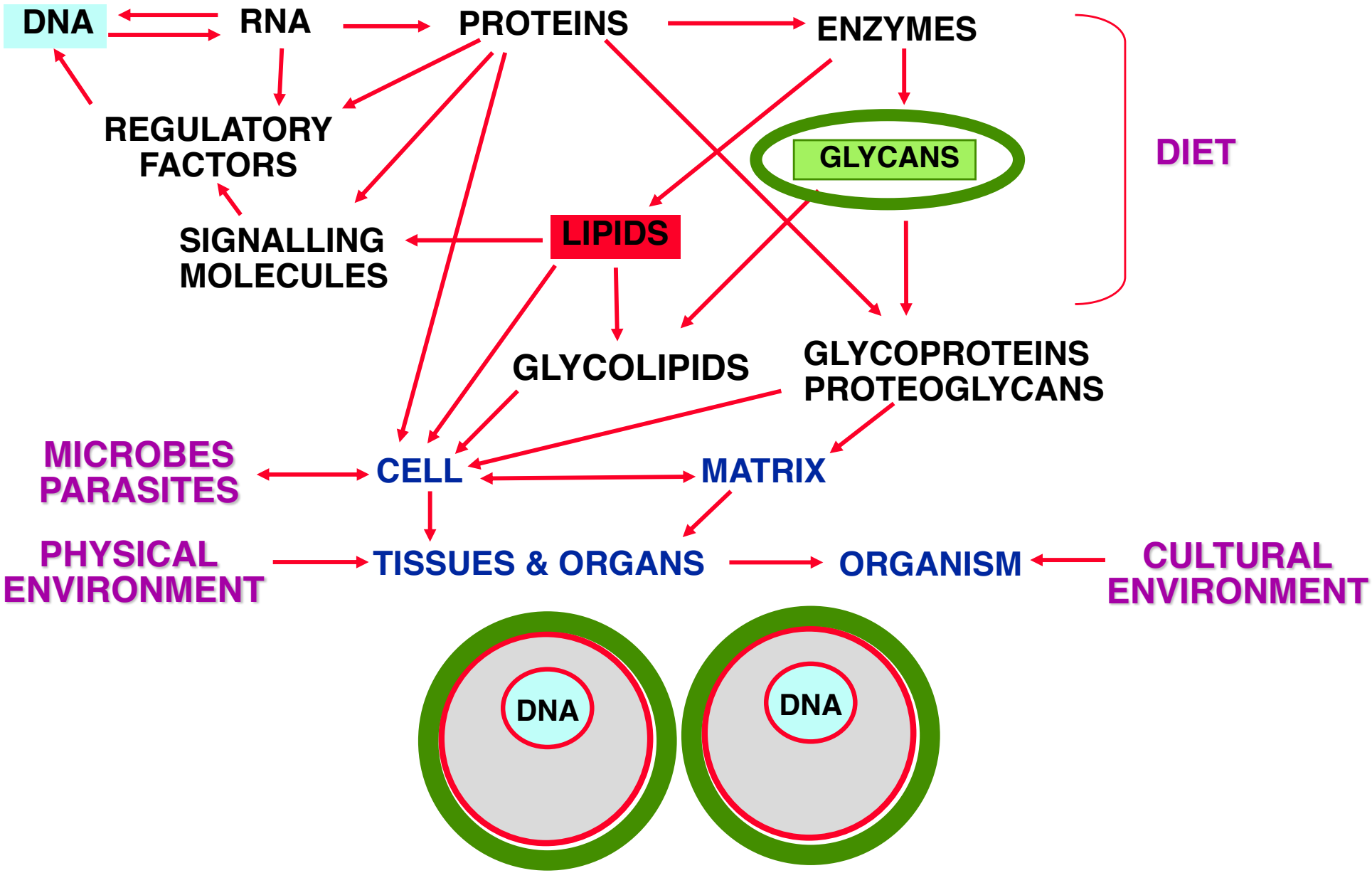
Extrinsic functions resulting from glycan–lectin interactions

DNA-Centric View of Molecular & Cellular Biology

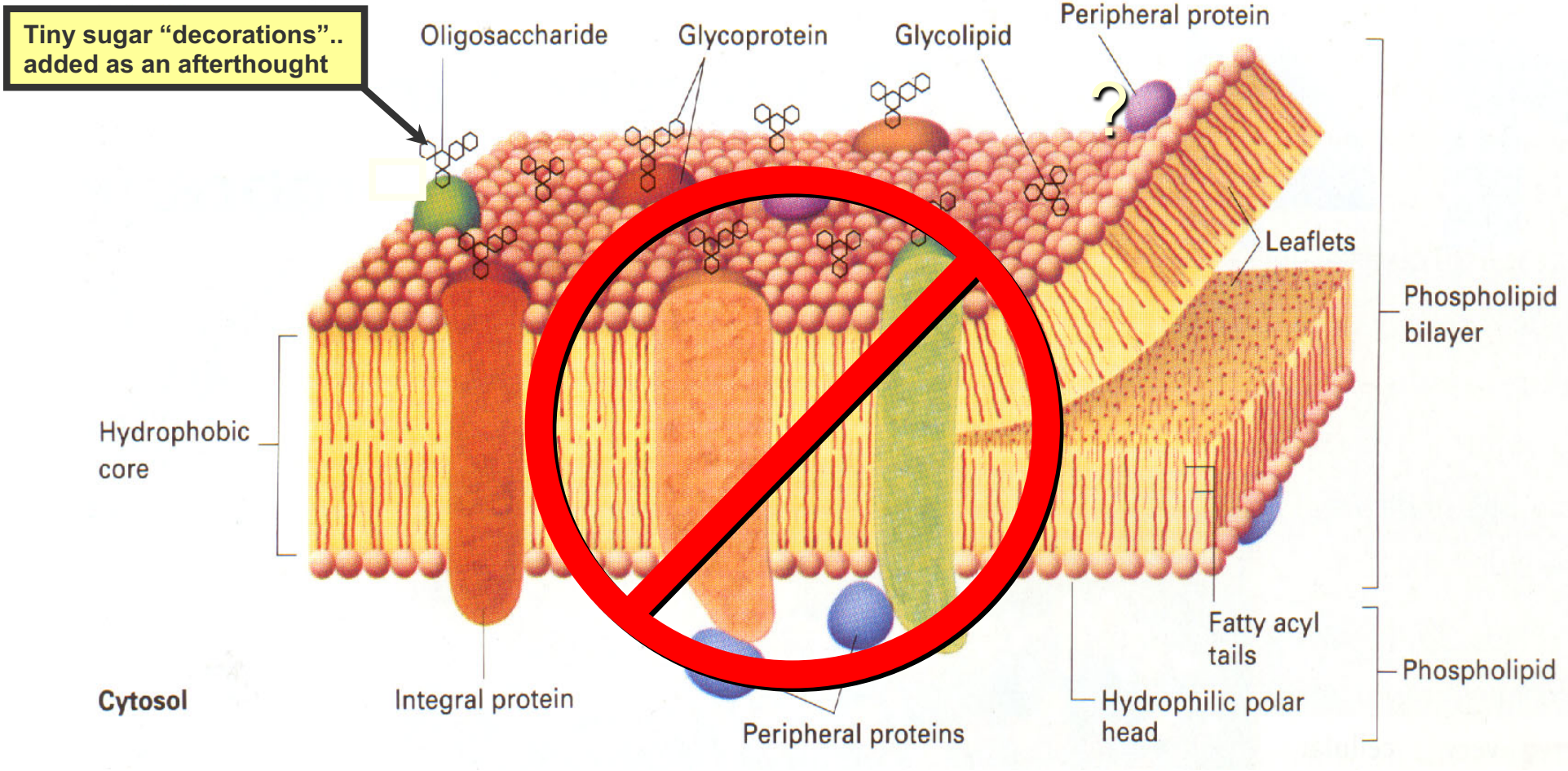
DNA → RNA → PROTEIN ⇢ CELL ⇢ ORGANISM



Holistic View of Molecular & Cellular Biology

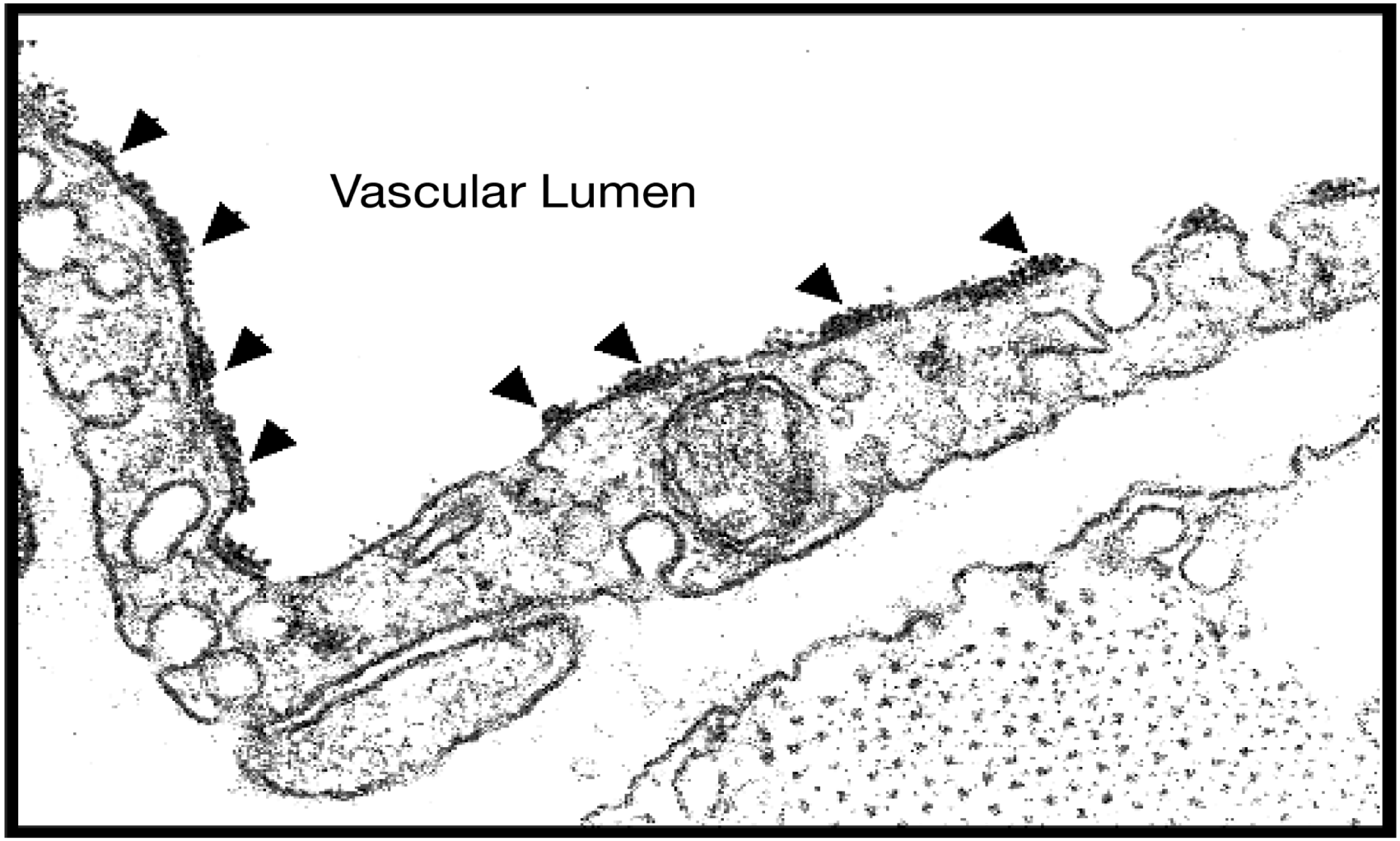


Eukaryotic Cell Surface circa 1995

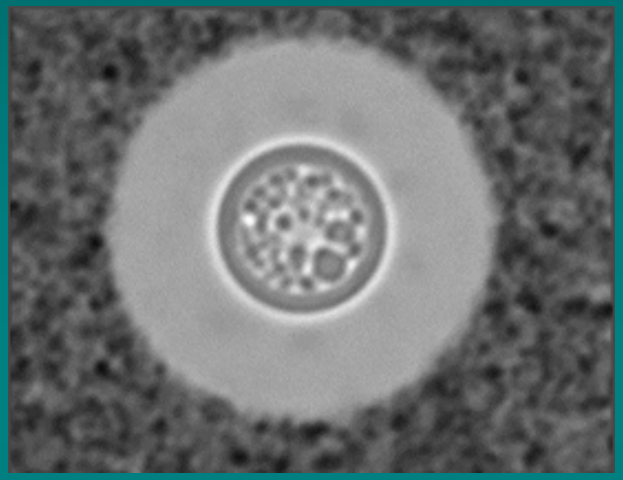


Molecular Cell Biology, 3rd Ed., Lodish, et al. (1995)

Historical electron micrograph of endothelial cells from a blood capillary in diaphragm muscle of a rat, showing the lumenal cell membrane of the cells (facing the blood) decorated with particles of cationized ferritin (*arrowheads*).



What if a Major Cell Component is Invisible by Light Microscopy?

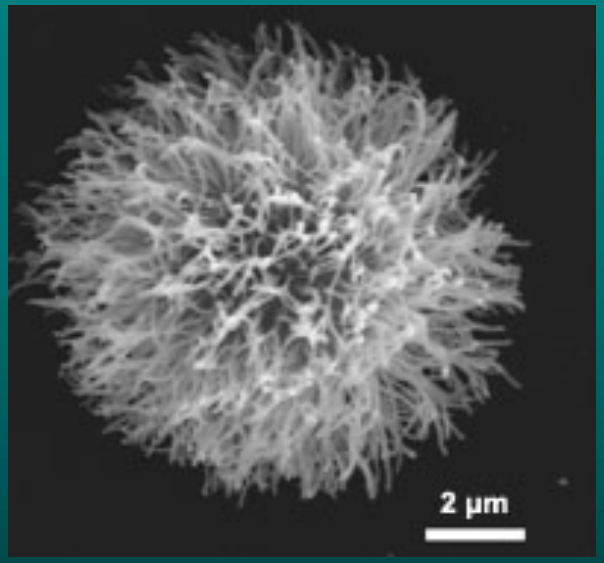


Light microscopy micrograph of *Cryptococcus neoformans* capsule delineated by India ink. The inner circle represents the fungal cell, with the wide outer circle being the capsule.

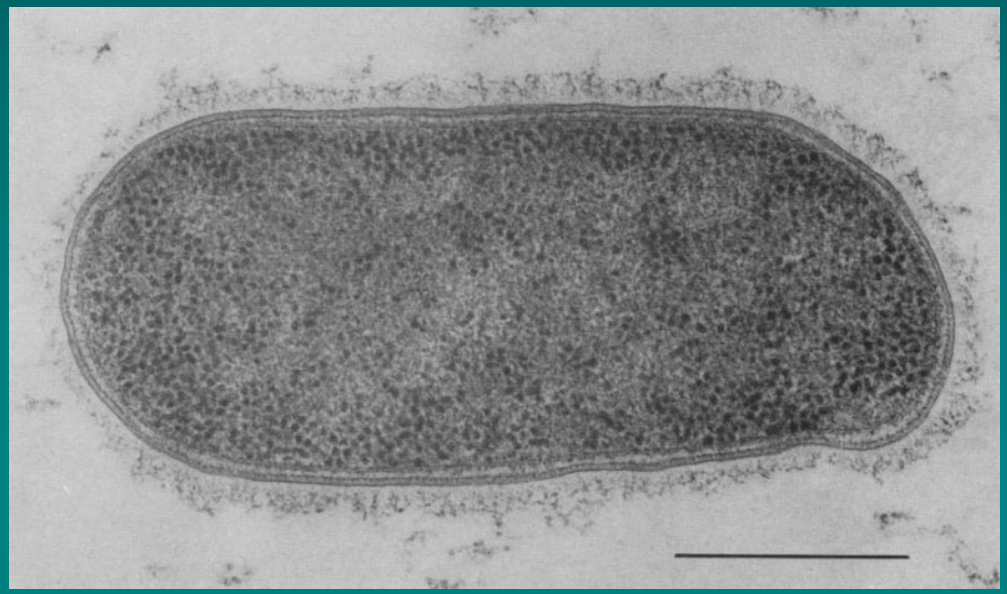
Steenbergen et al. (2003) *Microbes and Infection* 5:667

Scanning electron microscopy of *C. neoformans* yeast cells.

Van Duin et al. (2004) *Antimicrobial Agents and Chemotherapy* 48:2014

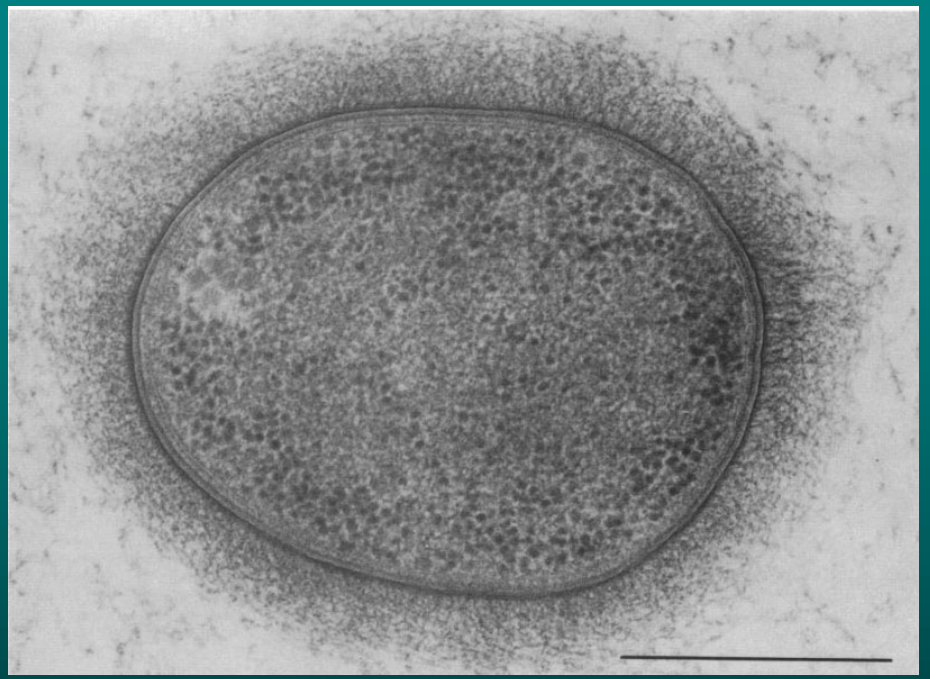


Unique to Yeast – NO!



Electron microscopic thin section of *Escherichia coli* K1

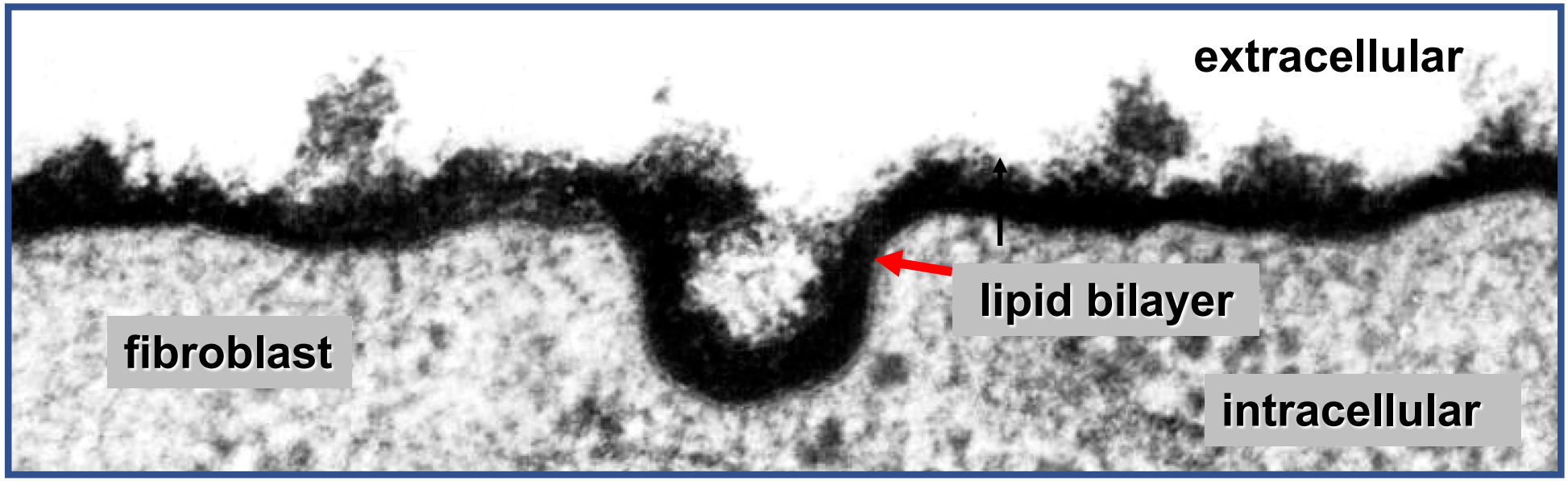
Amako et al. (1988) *J Bacteriol* 170:4960



Electron microscopic thin section of *Klebsiella pneumoniae*

Amako et al. (1988) *J Bacteriol* 170:4960

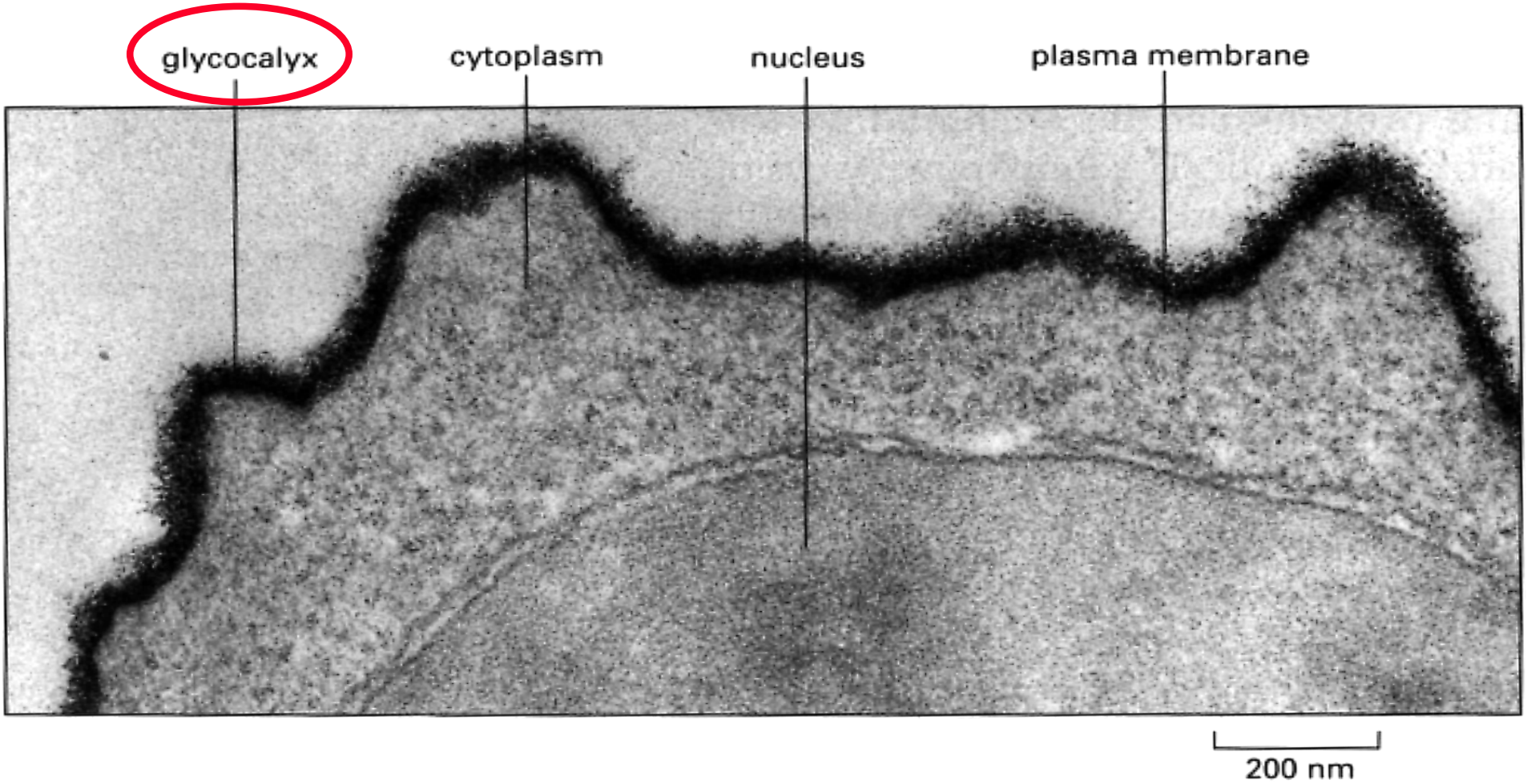
The Cell Surface – The Real Picture



Martinez-Palomo, A., *et al*. Cancer Res. 29, 925-937, 1969

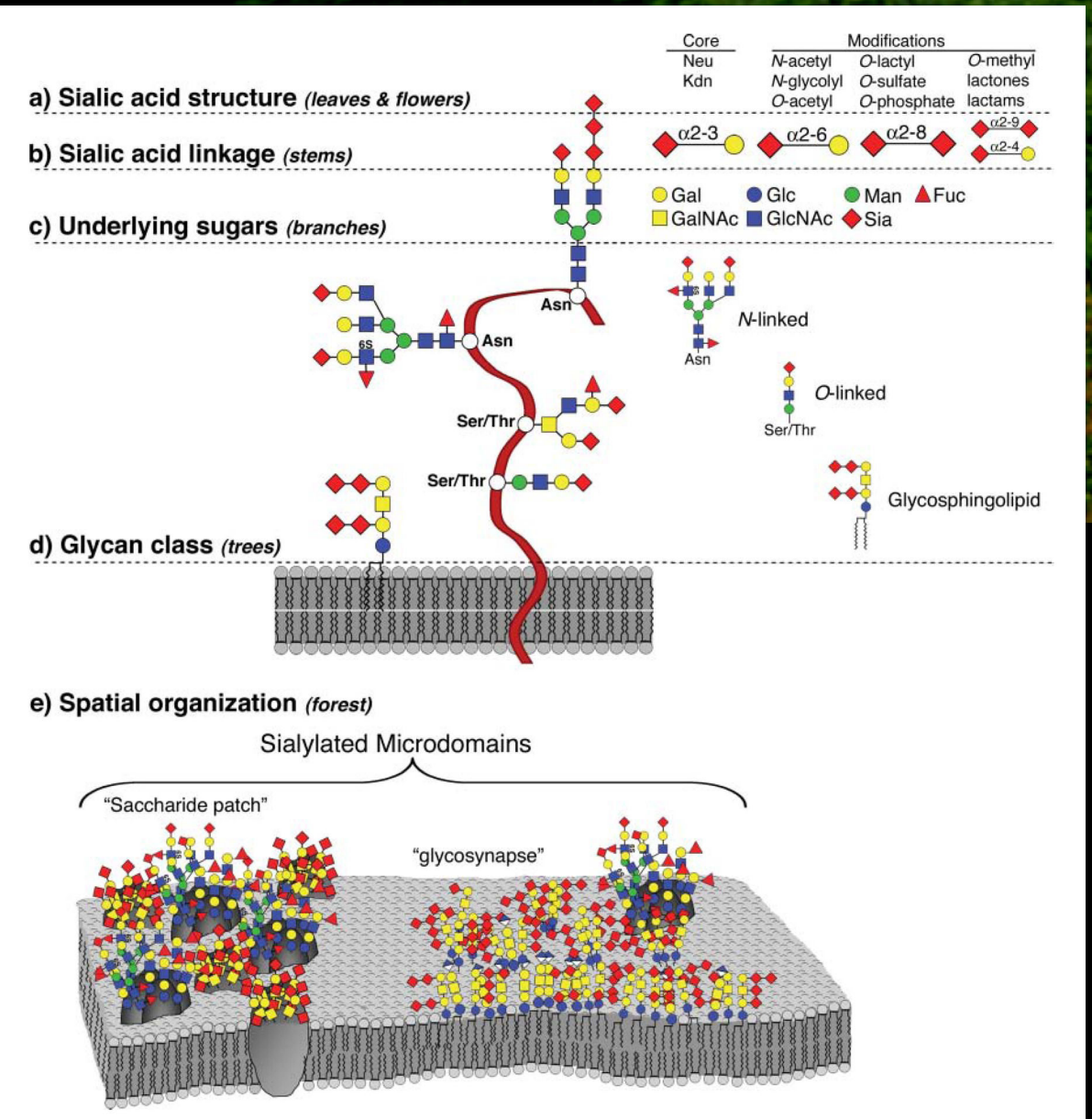
The *"glycocalyx"* surrounding a fibroblast. Cell surface carbohydrates are stained black.

All Cells Are Coated with Glycans



Electron micrograph of a human lymphocyte (Ruthenium Red staining)

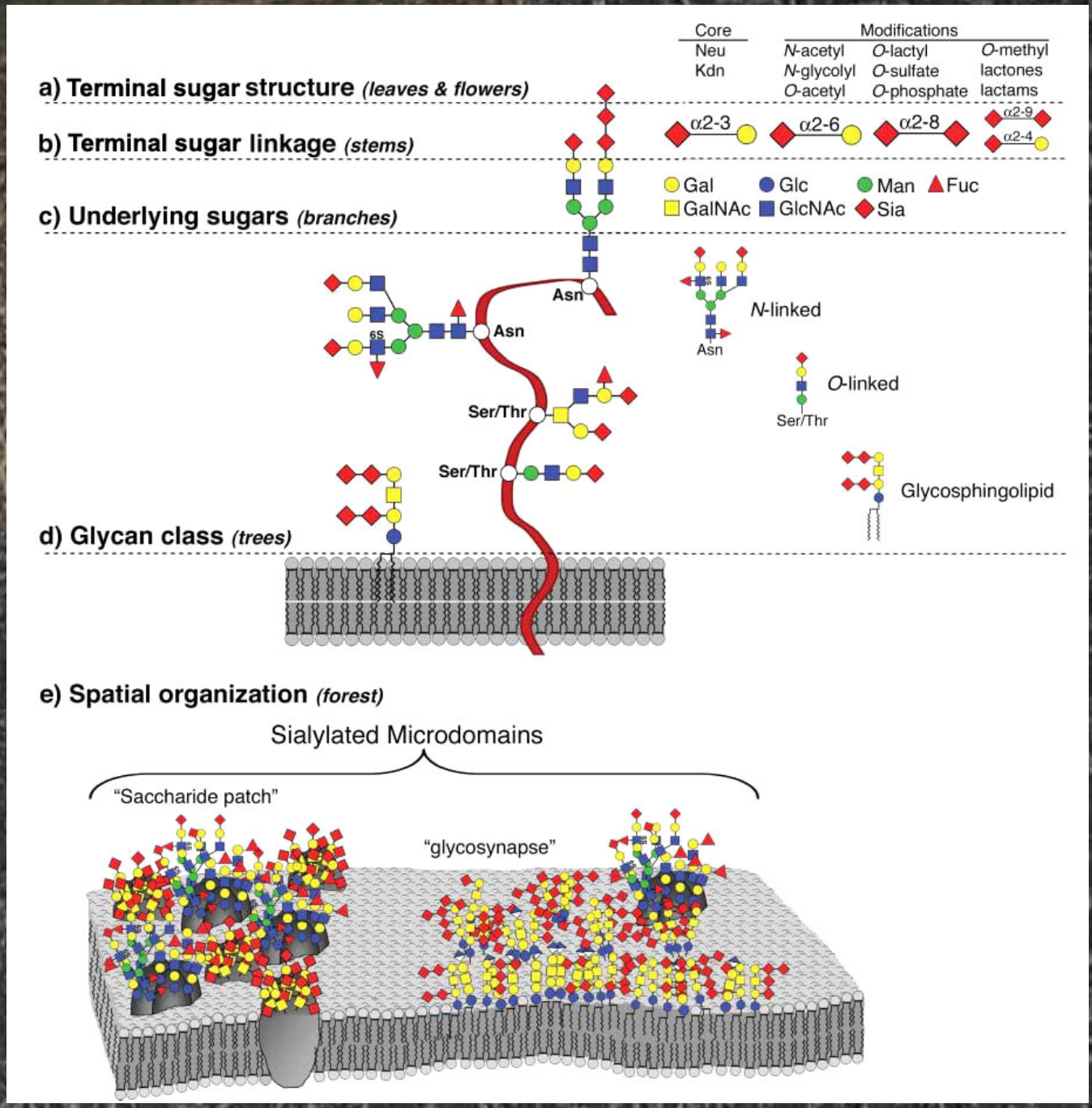
"Evolution has failed to generate a living cell devoid of surface glycosylation" - A. Varki



Tropical Forest Canopy

Cohen & Varki (2010)
OMICS 4:455

Glycobiology as a Language (Semiotics)

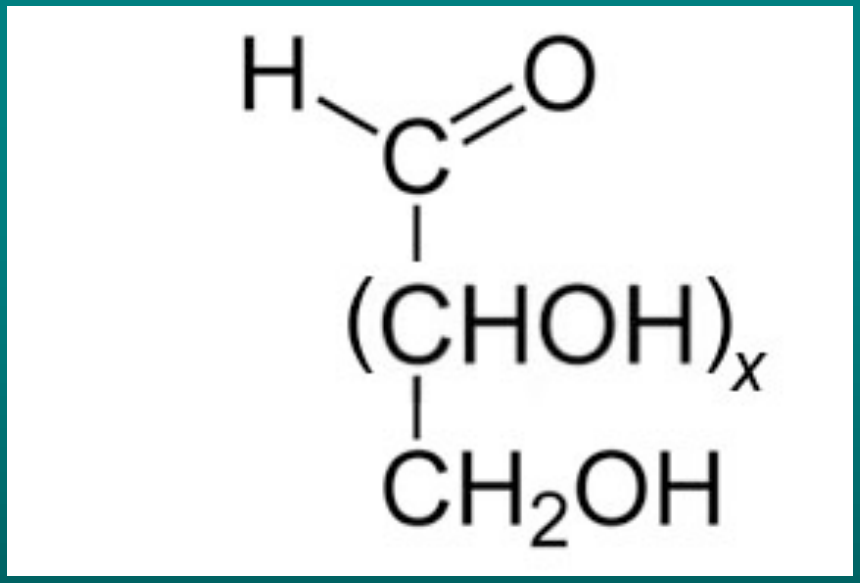


Cohen & Varki (2010)
OMICS 4:455

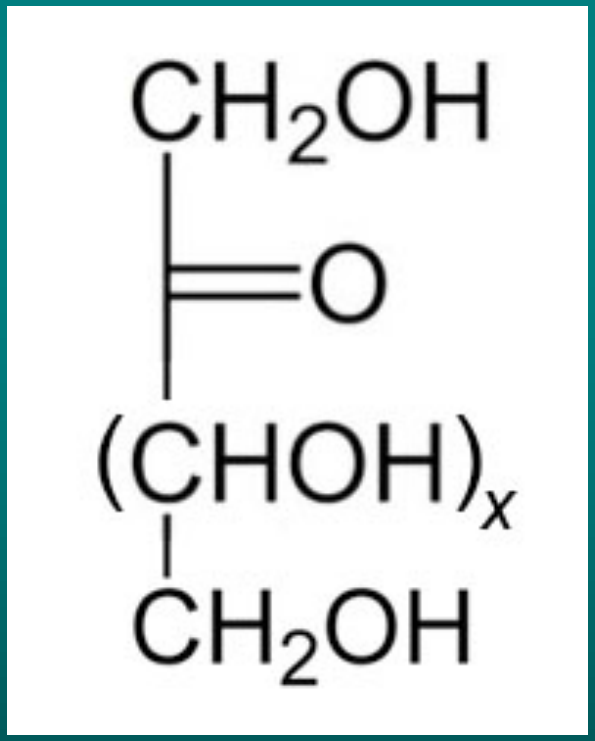
Rosetta Stone, British Museum

Carbohydrates – The Building Blocks of Glycobiology

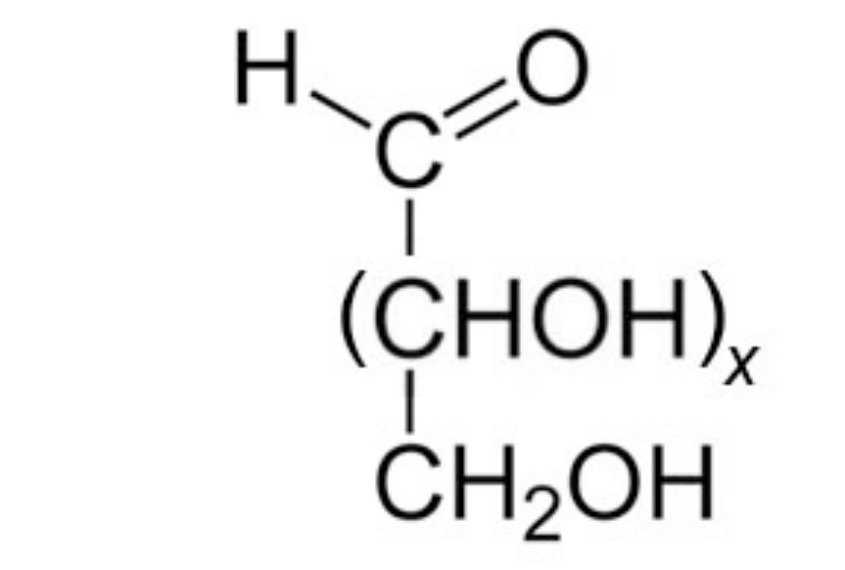
$(CH_2O)_n$ = "carbo" "hydrate"



Generalized Aldose



Generalized Ketose



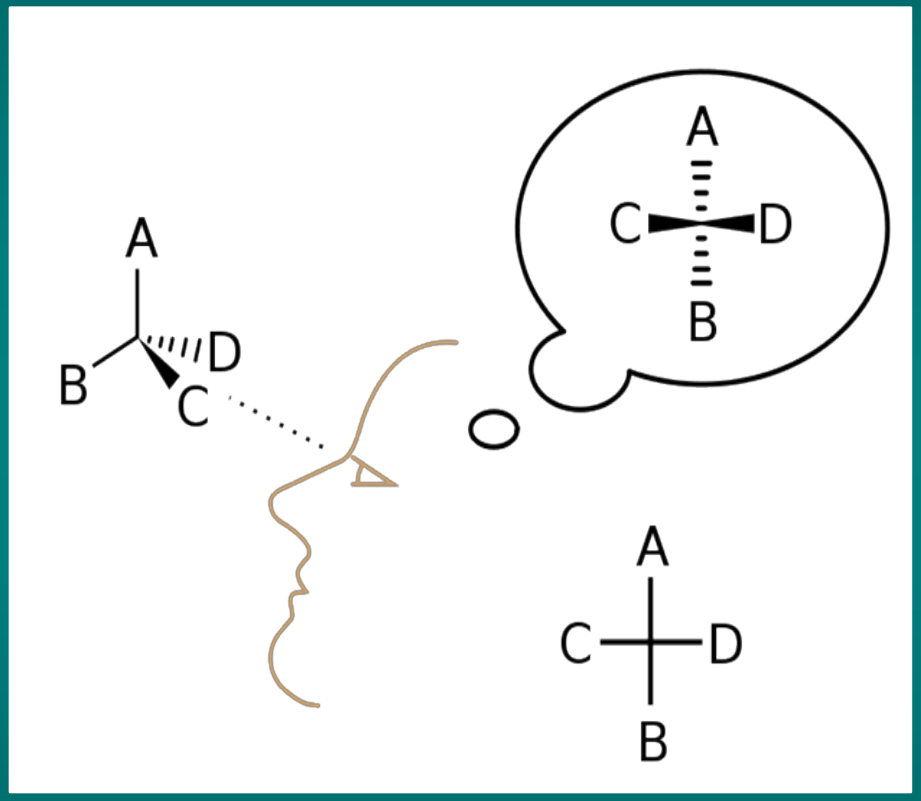
Generalized Aldose

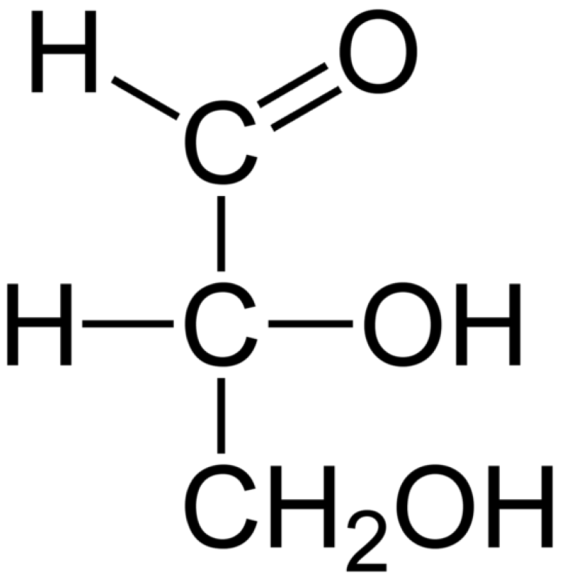
Monosaccharide identity is all about stereochemistry

Enantiomers:
Mirror images of each other that are not superimposable

Diastereomers:
Stereoisomers that are not enantiomers

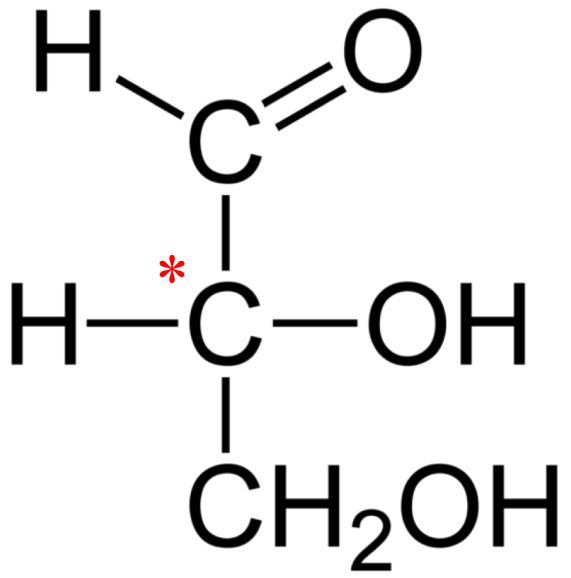
Fischer Projections



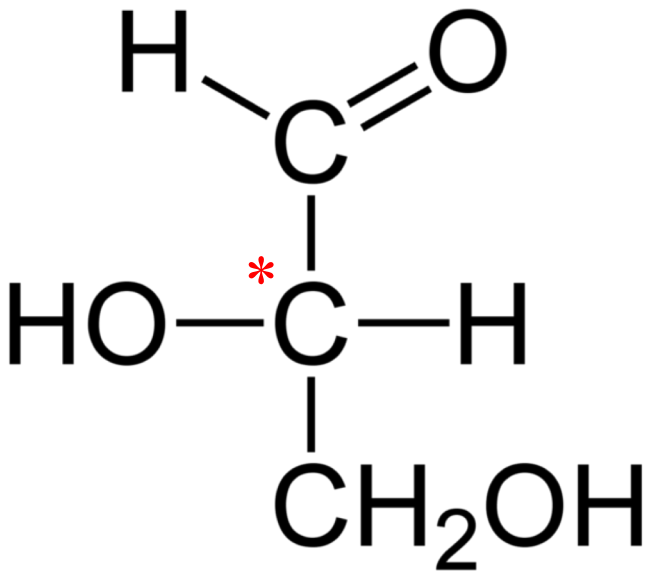


D-glyceraldehyde

Enantiomers



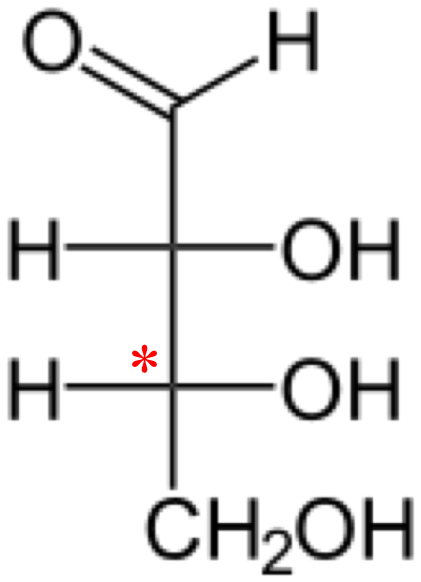
D-glyceraldehyde



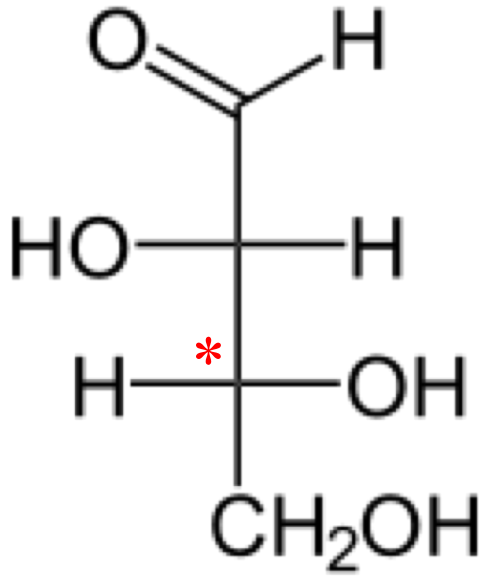
L-glyceraldehyde

* Highest numbered asymmetric carbon = reference carbon

Diasteriomers



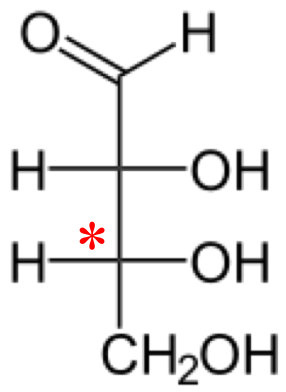
D-Erythrose



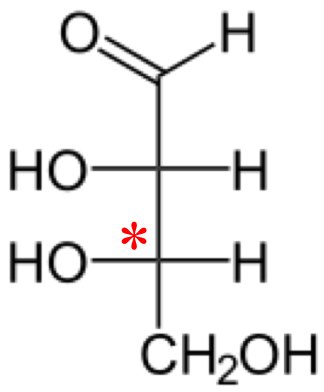
D-Threose

* Highest numbered asymmetric carbon = reference carbon

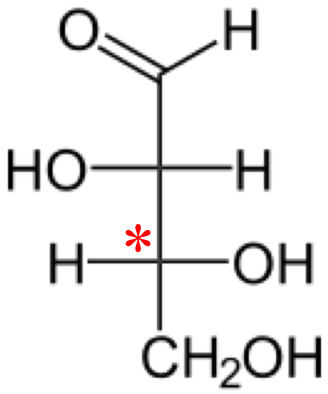
Enantiomers & Diasteriomers



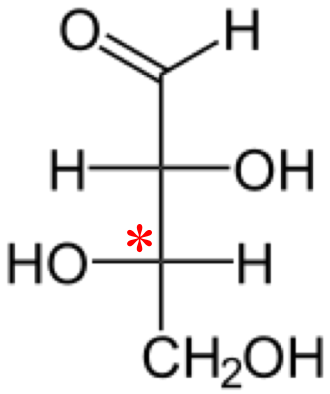
D-Erythrose



L-Erythrose



D-Threose

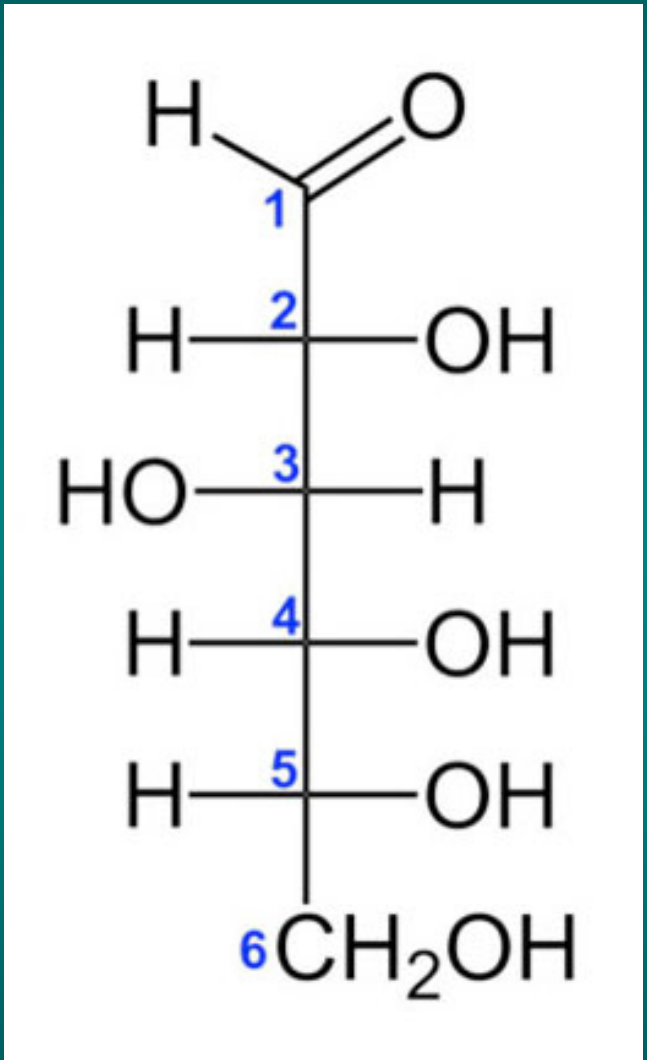


L-Threose

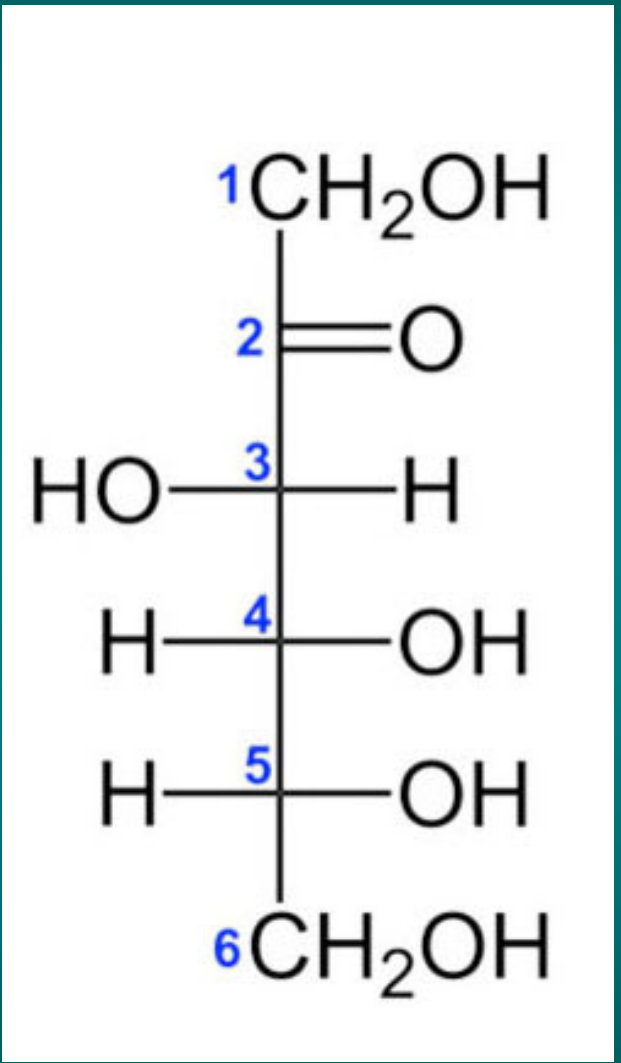
* Highest numbered asymmetric carbon = reference carbon

# carbon	asymmetric carbons	diaster-iomers	diasteriomers & enantiomers
3	1	0	2
4	2	2	4
5	3	4	8
6	4	8	16

Example of Aldose & Ketose

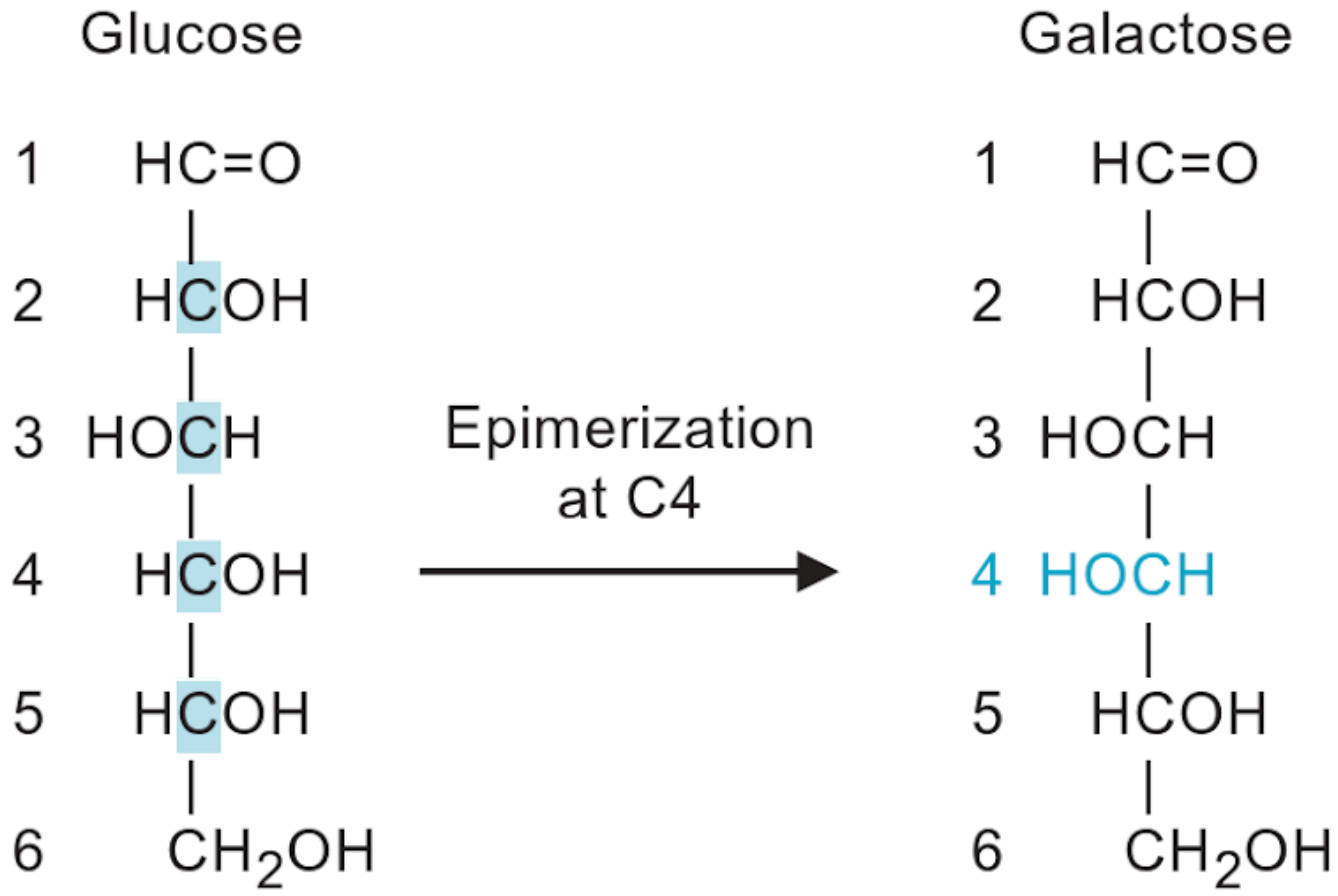


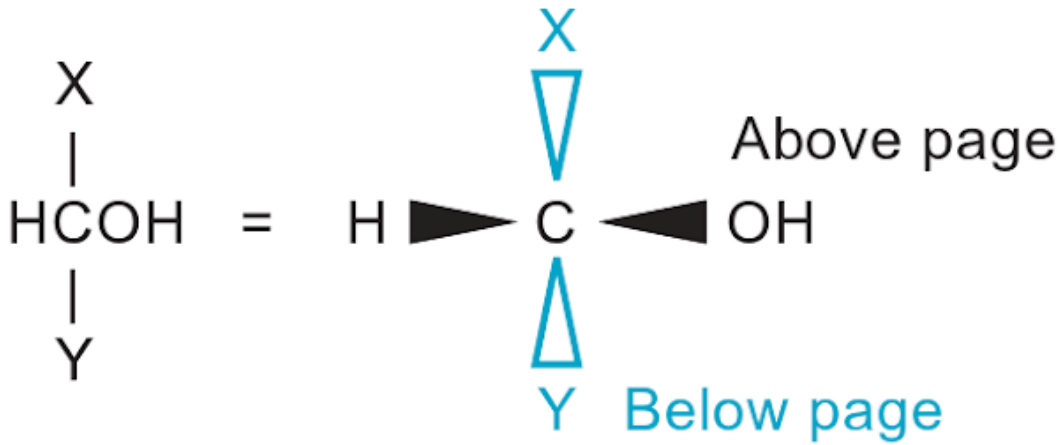
Glucose



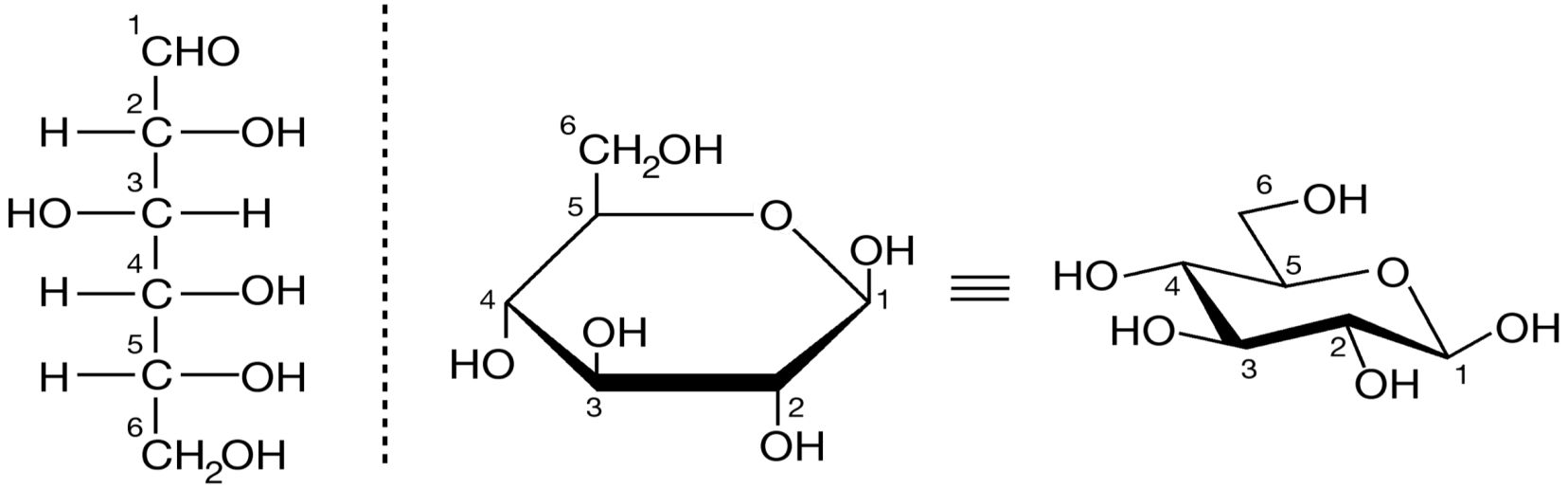
Fructose

Stereochemistry of Glucose & Galactose





Monosaccharides – The Basic Structural Unit

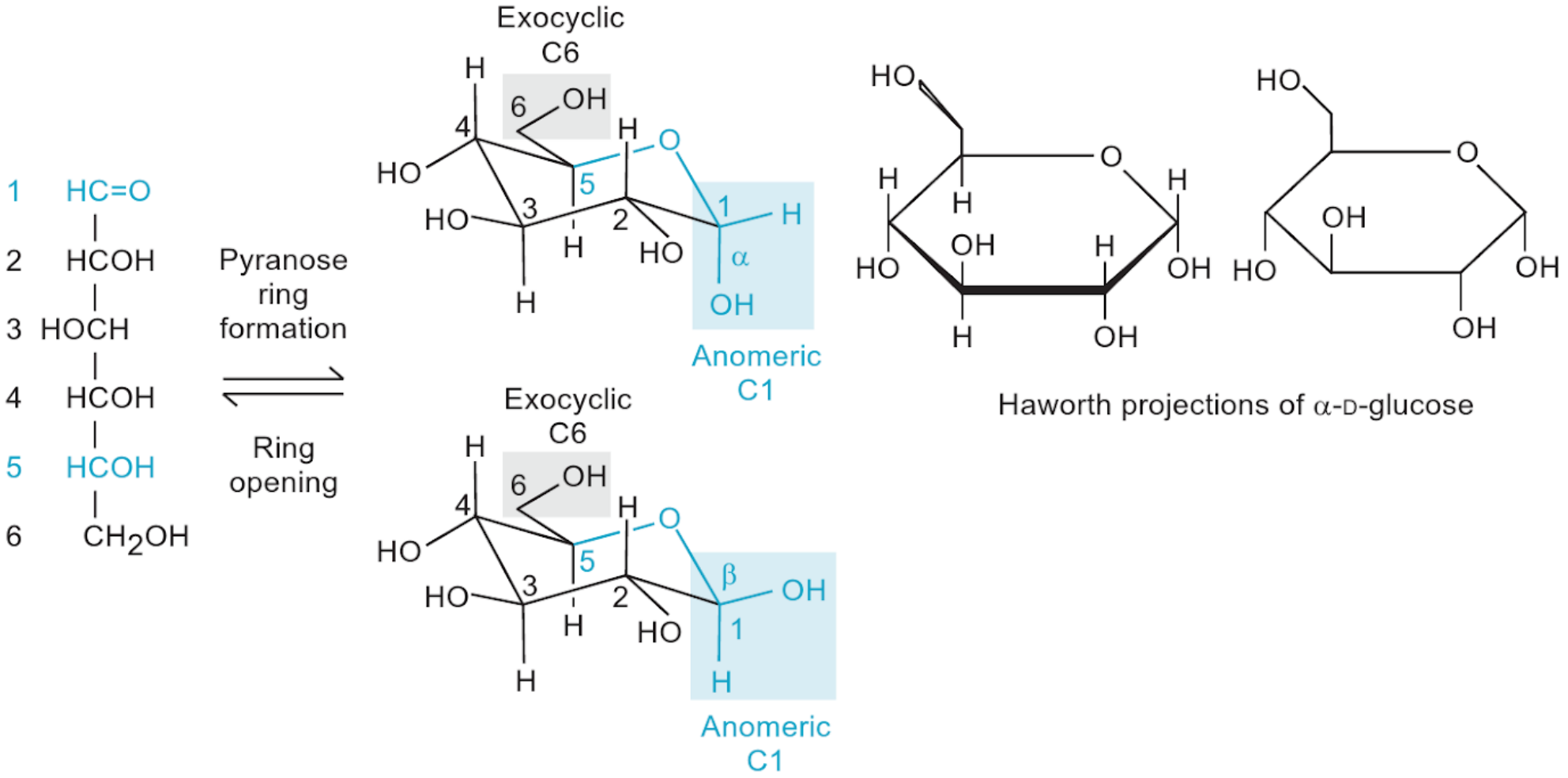


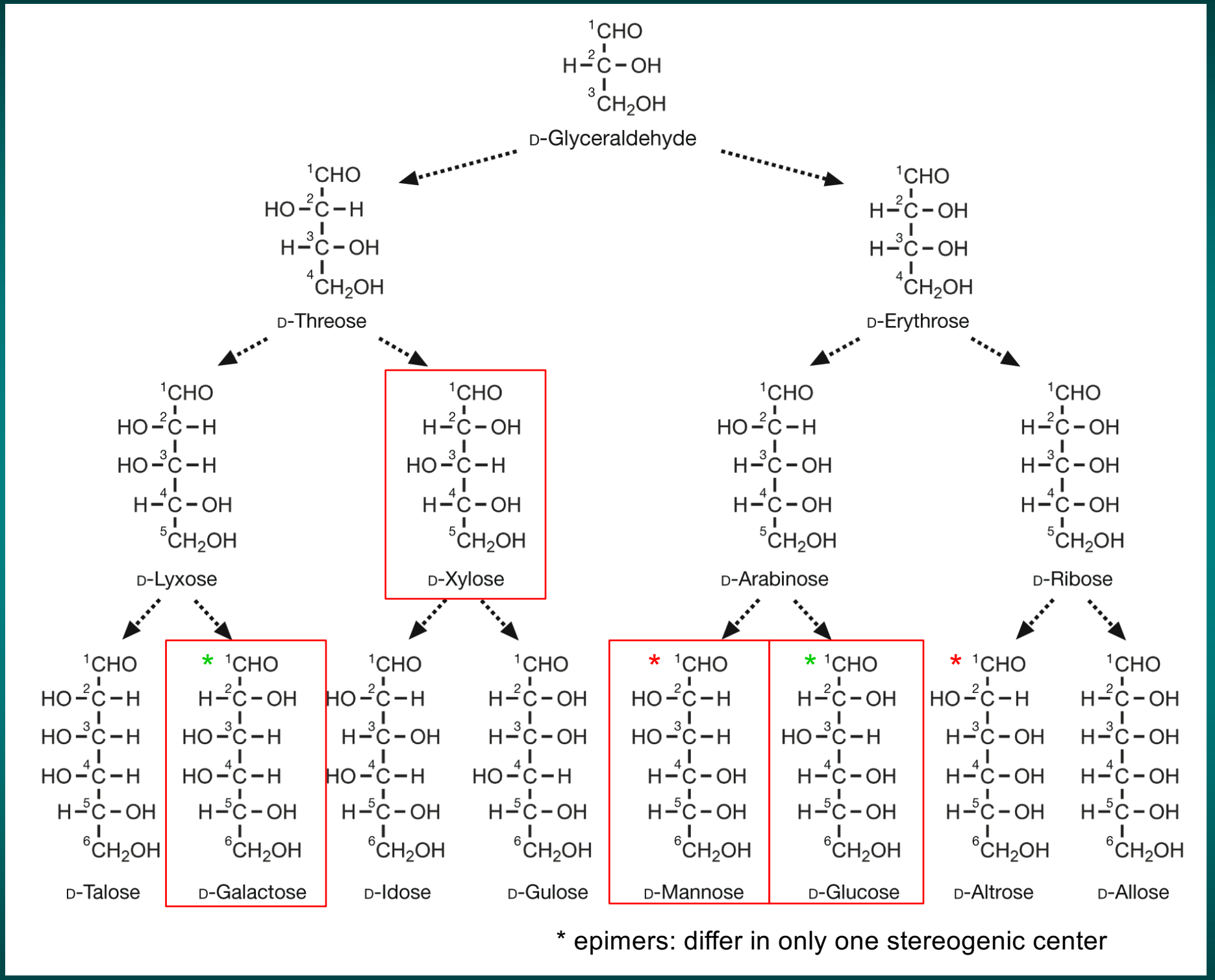
D-glucose

β-D-glucose

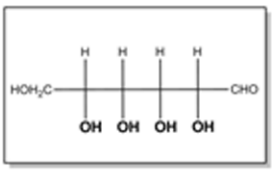
- **Carbonyl group at the end of the carbon chain (aldoses) or at an inner carbon (ketoses) has potential reducing power. This end is called the reducing terminus, or reducing end**
- **The ring form of a monosaccharide generates a chiral (anomeric) center (at C-1 for aldo sugars or at C-2 for keto sugars). Notice that other positions are chiral, which therefore imparts stereochemical information**

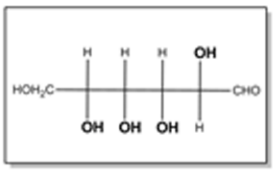
Structural Representations of Glucose: Fischer & Haworth Projections, Anomeric Center

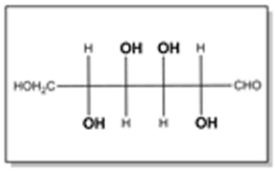


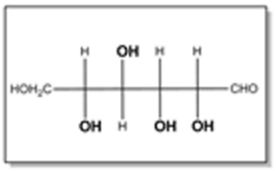

D-Glyceraldehyde
D-Threose
D-Erythrose
D-Lyxose
D-Xylose
D-Arabinose
D-Ribose
D-Talose
D-Galactose
D-Idose
D-Gulose
D-Mannose
D-Glucose
D-Altrose
D-Allose
* epimers: differ in only one stereogenic center

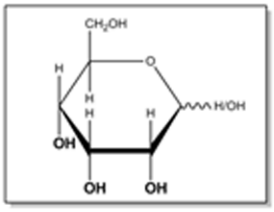
Graphical Depiction of Monosaccharides

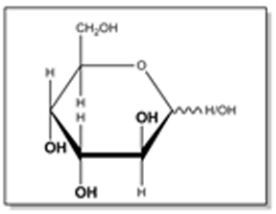


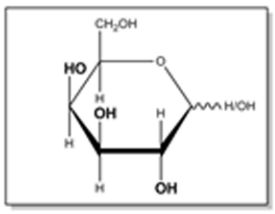


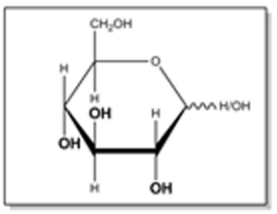


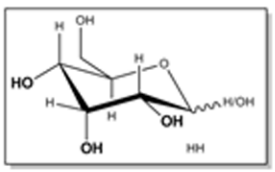


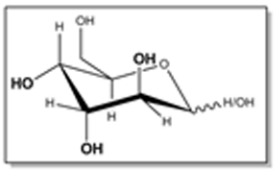


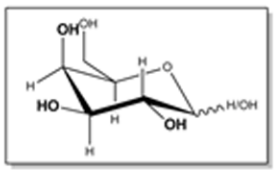


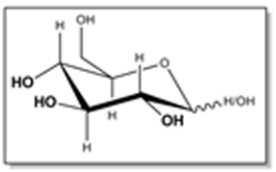


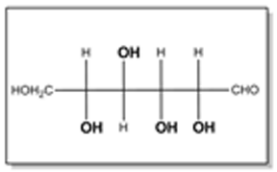


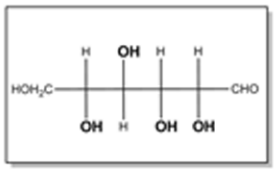


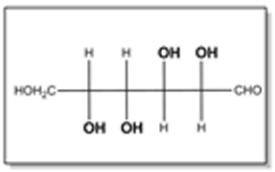


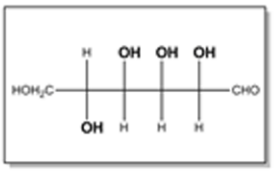


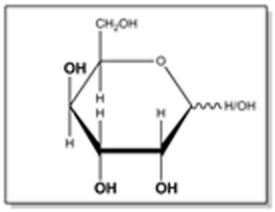


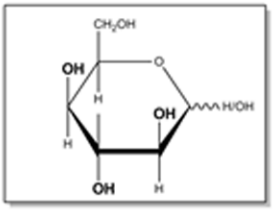


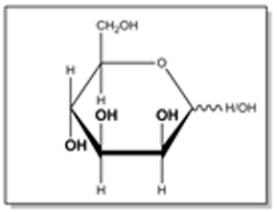


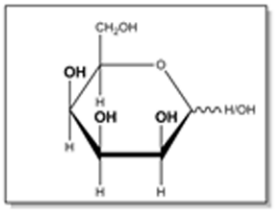


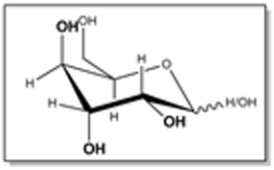


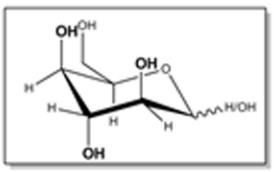


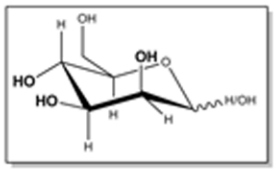


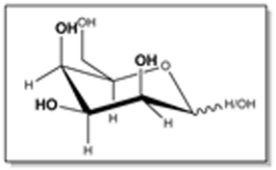


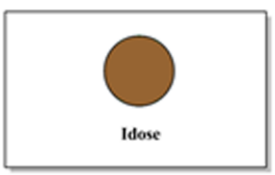


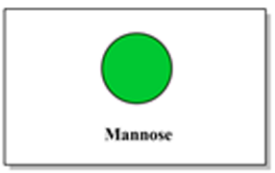


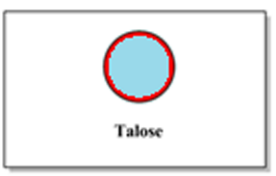




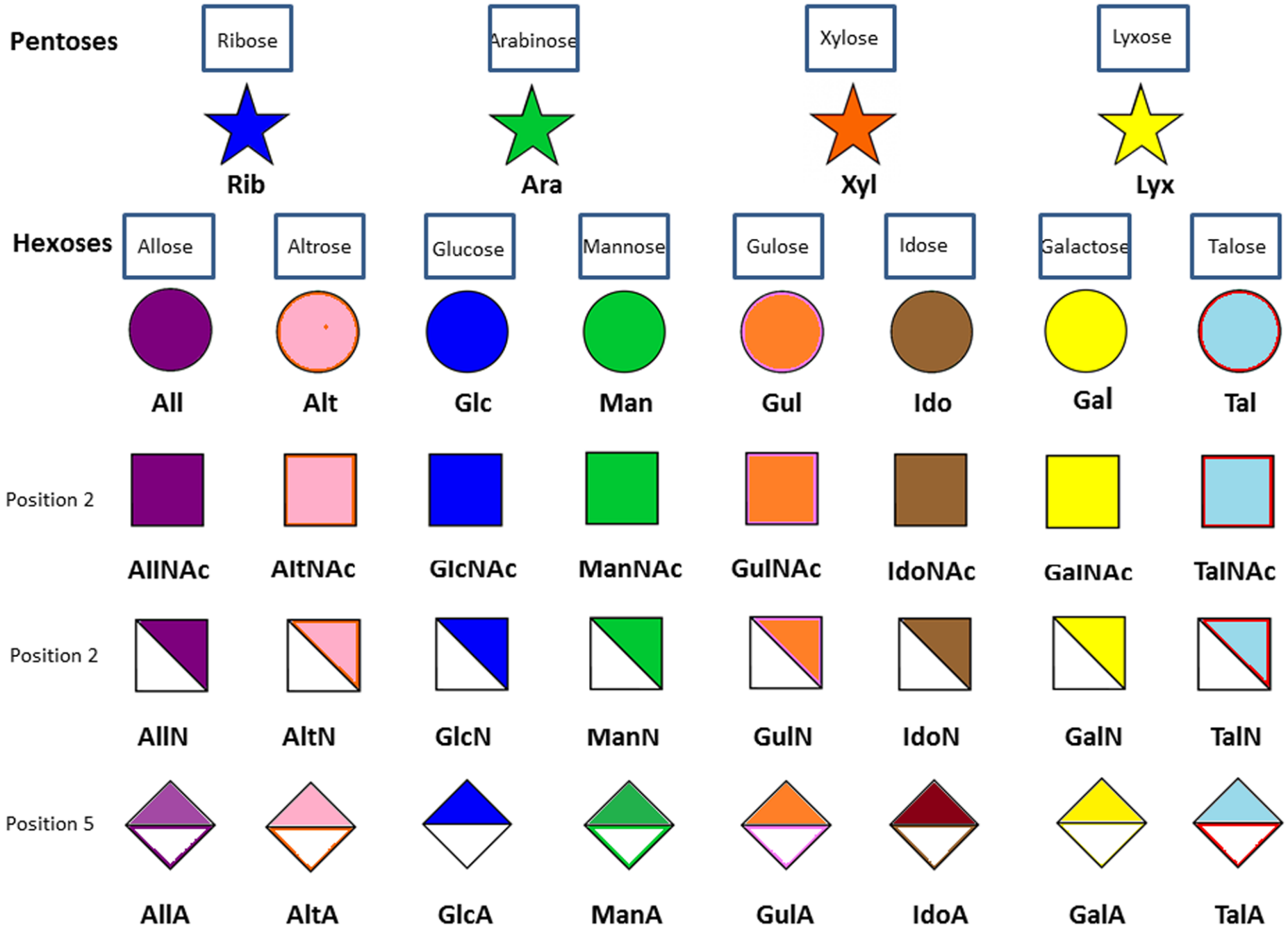








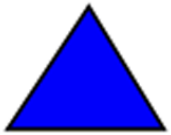
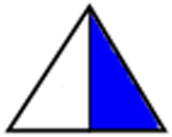
Monosaccharide Nomenclature Symbols



More Monosaccharide Symbols

6-Deoxy Sugars

NAc and Amines are in the 2 position

6-deoxy-D-Galactose (D-Fucose)	6-deoxy-D Glucose (D-Quinvose)	6-deoxy-D-Mannose (D-Rhamnose)
		
Fuc	Qui	Rha
		
FucNAc	QuiNAc	RhaNAc
		
FucN	QuiN	RhaN

Sialic Acids


Neu5Ac


Neu5Gc


KDN

Ketoses


Fru


Sor


Psi


Tag

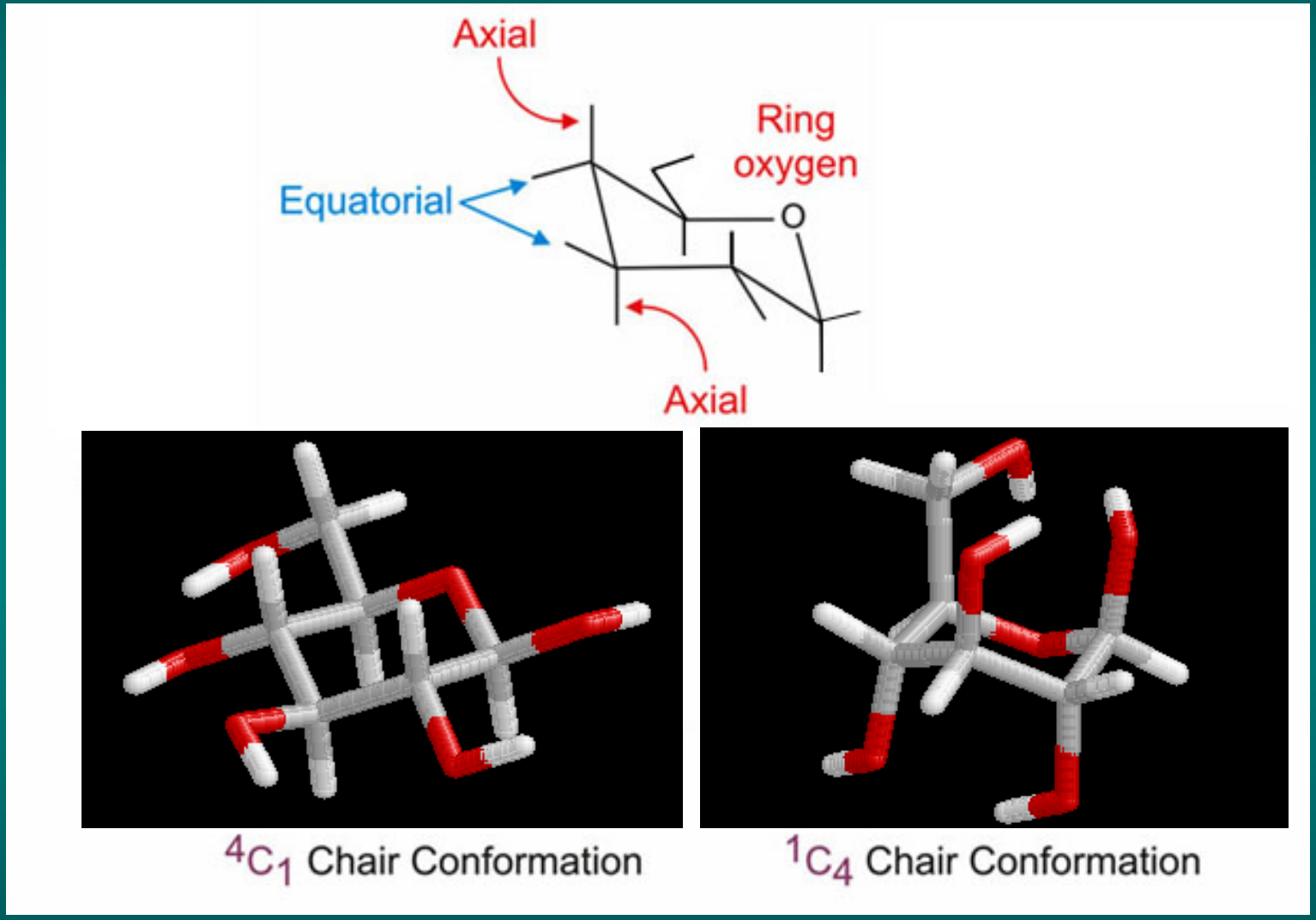
Essentials of Glycobiology, 3rd Edition — Symbol Nomenclature for Glycans (Appendix 1B)

Reference: Symbol Nomenclature for Graphical Representation of Glycans, *Glycobiology,* ***25***, 1323-1324, 2015.

Monosaccharide Symbol Nomenclature:

SHAPE	White	Blue	Green	Yellow	Orange	Pink	Purple	Light Blue	Brown	Red
Filled Circle	Hexose	Glc	Man	Gal	Gul	Alt	All	Tal	Ido	
Filled Square	HexNAc	GlcNAc	ManNAc	GalNAc	GulNAc	AltNAc	AllNAc	TalNAc	IdoNAc	
Crossed Square	Hexosamine	GlcN	ManN	GalN	GulN	AltN	AllN	TalN	IdoN	
Divided Diamond	Hexuronate	GlcA	ManA	GalA	GulA	AltA	AllA	TalA	IdoA	
Filled Triangle	Deoxyhexose	Qui	Rha		6dGul	6dAlt		6dTal		Fuc
Divided Triangle	DeoxyhexNAc	QuiNAc	RhaNAc			6dAltNAc		6dTalNAc		FucNAc
Flat Rectangle	Dideoxyhexose	Oli	Tyv		Abe	Par	Dig	Col		
Filled Star	Pentose		Ara	Lyx	Xyl	Rib				
Filled Diamond	Deoxynonulosonate		Kdn				Neu5Ac	Neu5Gc	Neu	Sia
Flat Diamond	Dideoxynonulosonate		Pse	Leg		Aci		4eLeg		
Flat Hexagon	Unknown	Bac	LDmanHep	Kdo	Dha	DDmanHep	MurNAc	MurNGc	Mur	
Pentagon	Assigned	Api	Fru	Tag	Sor	Psi				

Monosaccharide Conformation – Chair Configurations



Taylor & Drickamer (2006) Introduction to Glycobiology, 2nd edition

Nearly all Vertebrate Glycans are Built from Only 9 Sugars

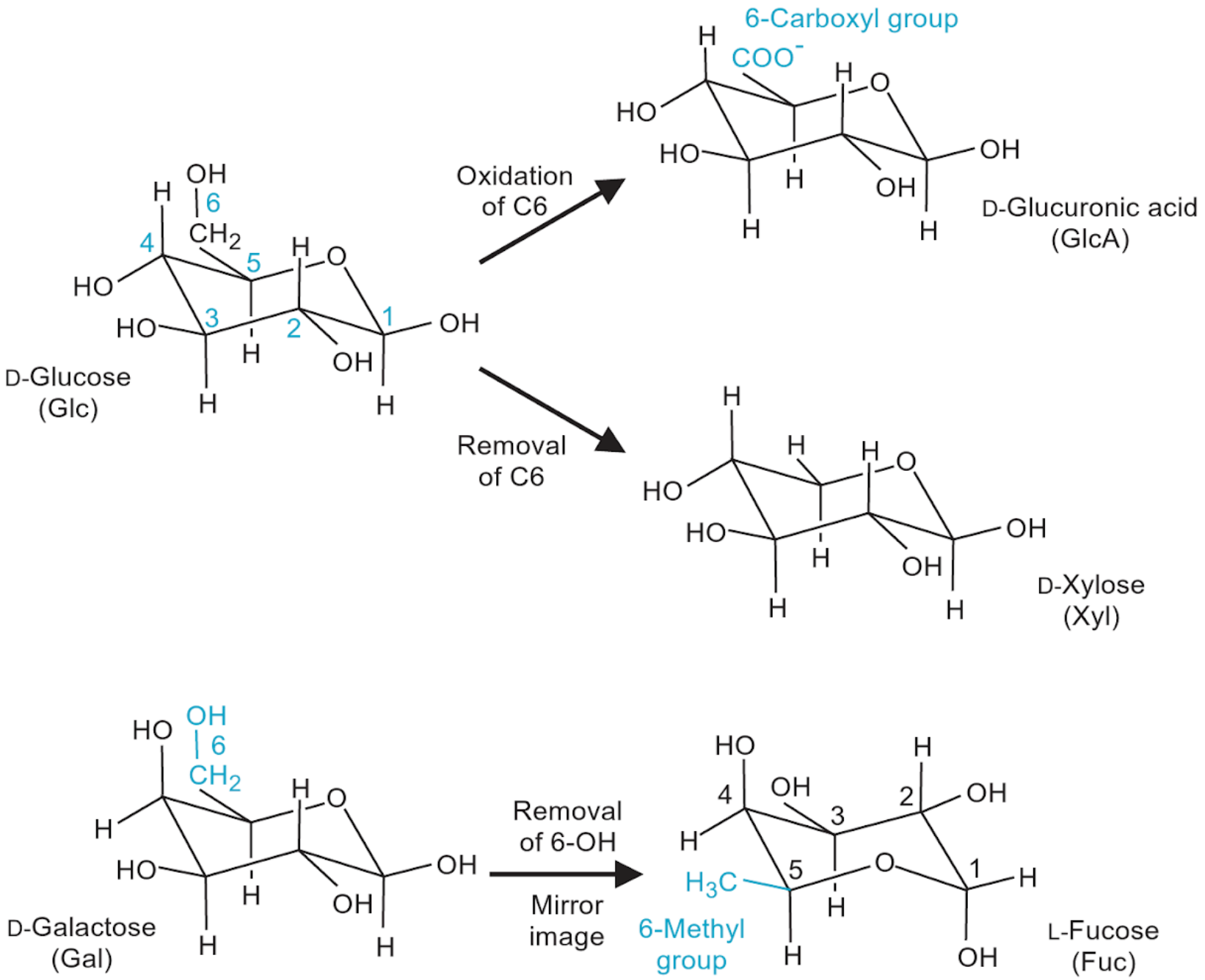
- 3 hexoses*
  - Glucose
  - Mannose
  - Galactose
- 2 N-acetylhexosamines
  - N-acetylglucosamine
  - N-acetylgalactosamine
- xylose
- glucuronic acid
- sialic acid
- L-fucose

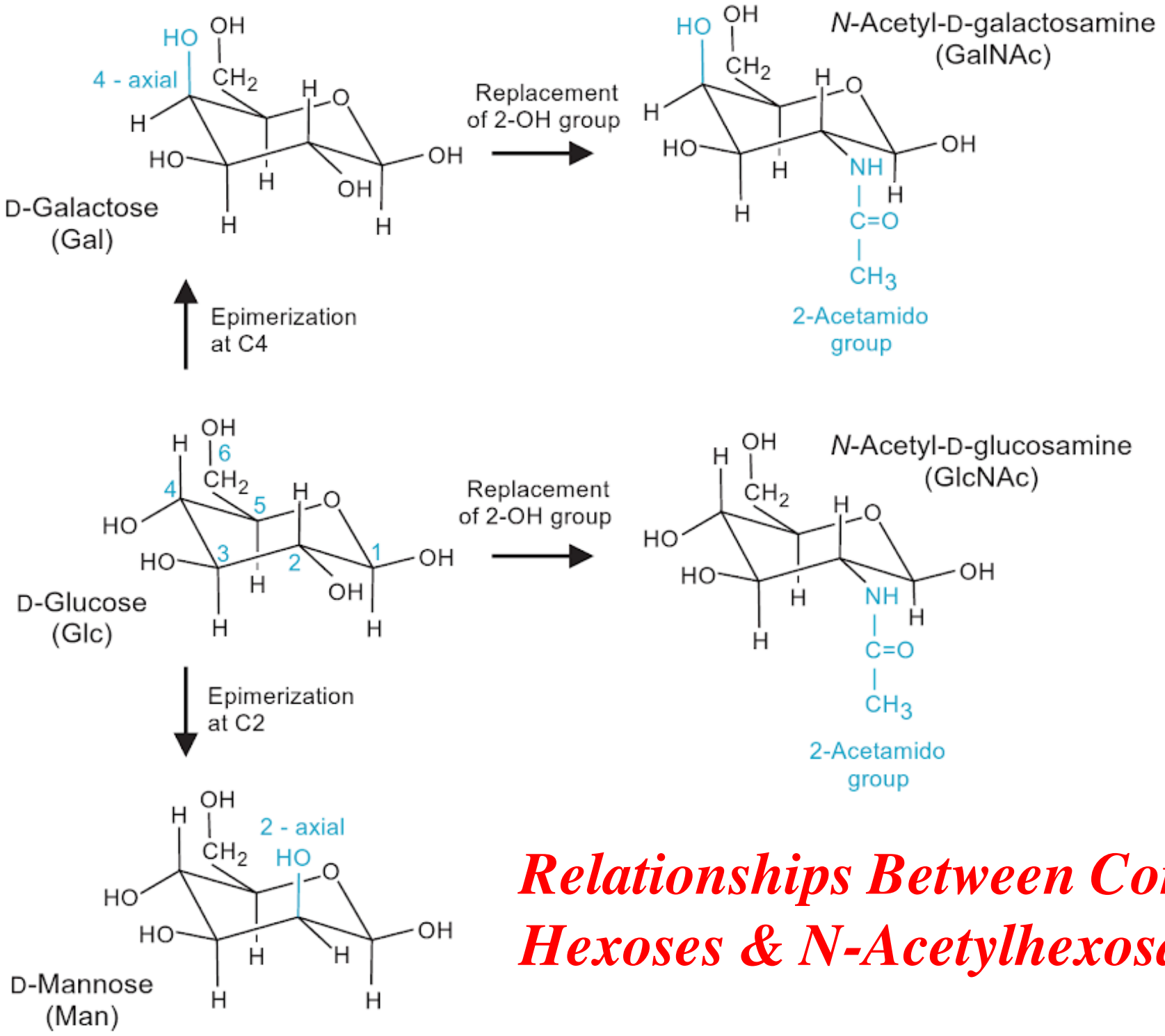
*all D configuration except fucose

Mammalian Monosaccharide Abundance

Monosaccharides	Abundance (%)
D-GlcNAc	31.8
D-Gal	24.8
D-Man	18.9
Neu5Ac	8.3
L-Fuc	7.2
D-GalNAc	4.8
D-Glc	2.5
D-GlcA	0.3
D-Xyl	0.1
L-IdoA	0.1
Others	1.2

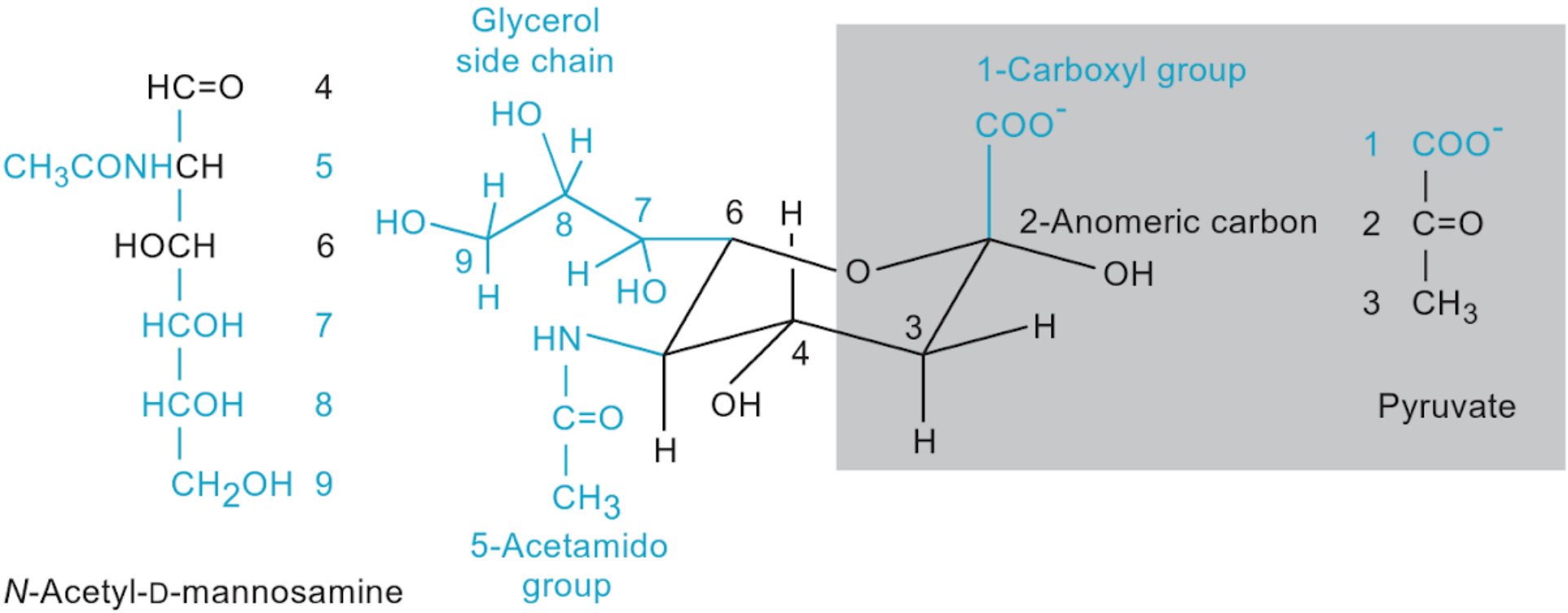
Common Derivatives of the Hexoses





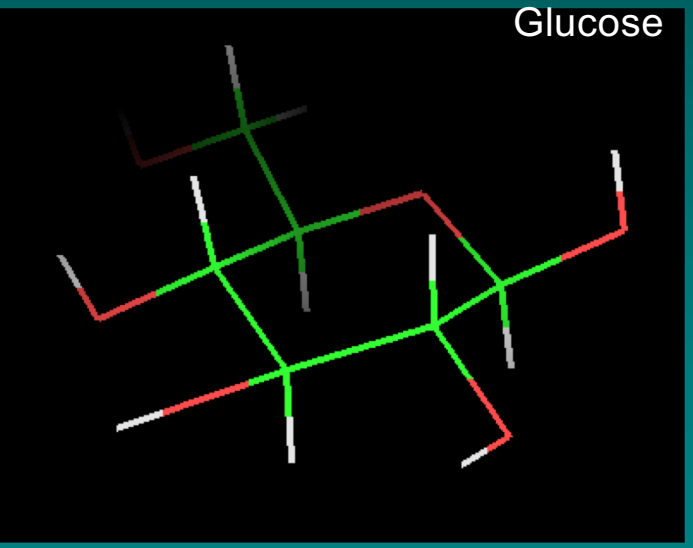
Relationships Between Common Hexoses & N-Acetylhexosamines

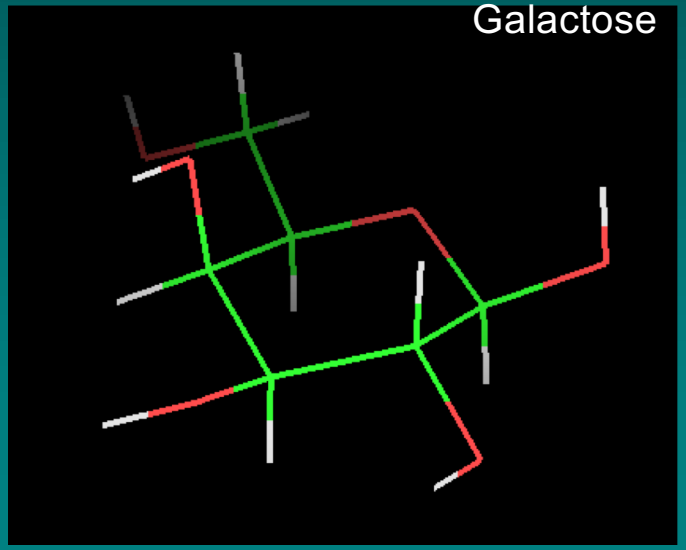
N-Acetylneuraminic Acid: The Most Common Sialic Acid

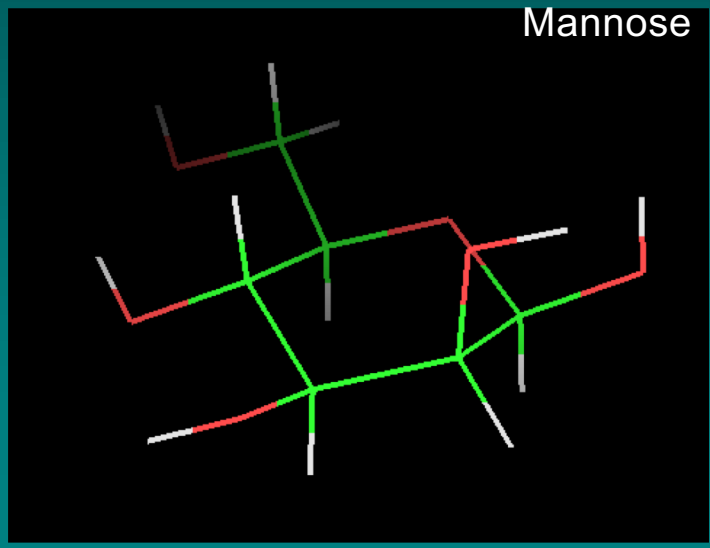


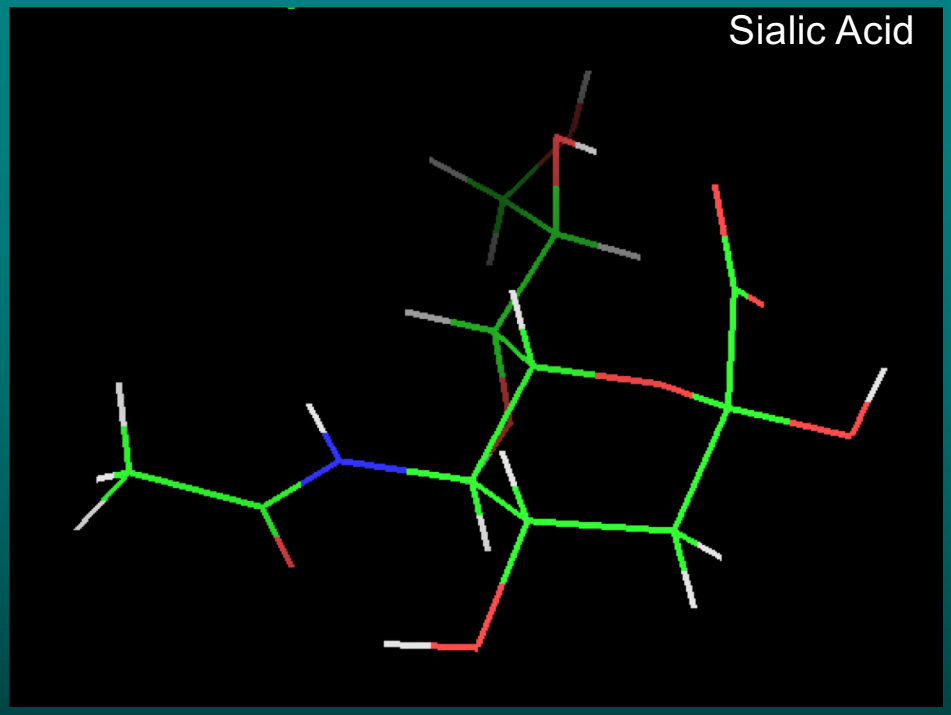
Carbons C1-C3 are derived from Pyruvate, and C4-C9 are from *N*-Acetylmannosamine

Glycan Properties for Molecular Recognition & Binding Energy









Seven Eukaryotic Sugars & Their Relationship to Glucose