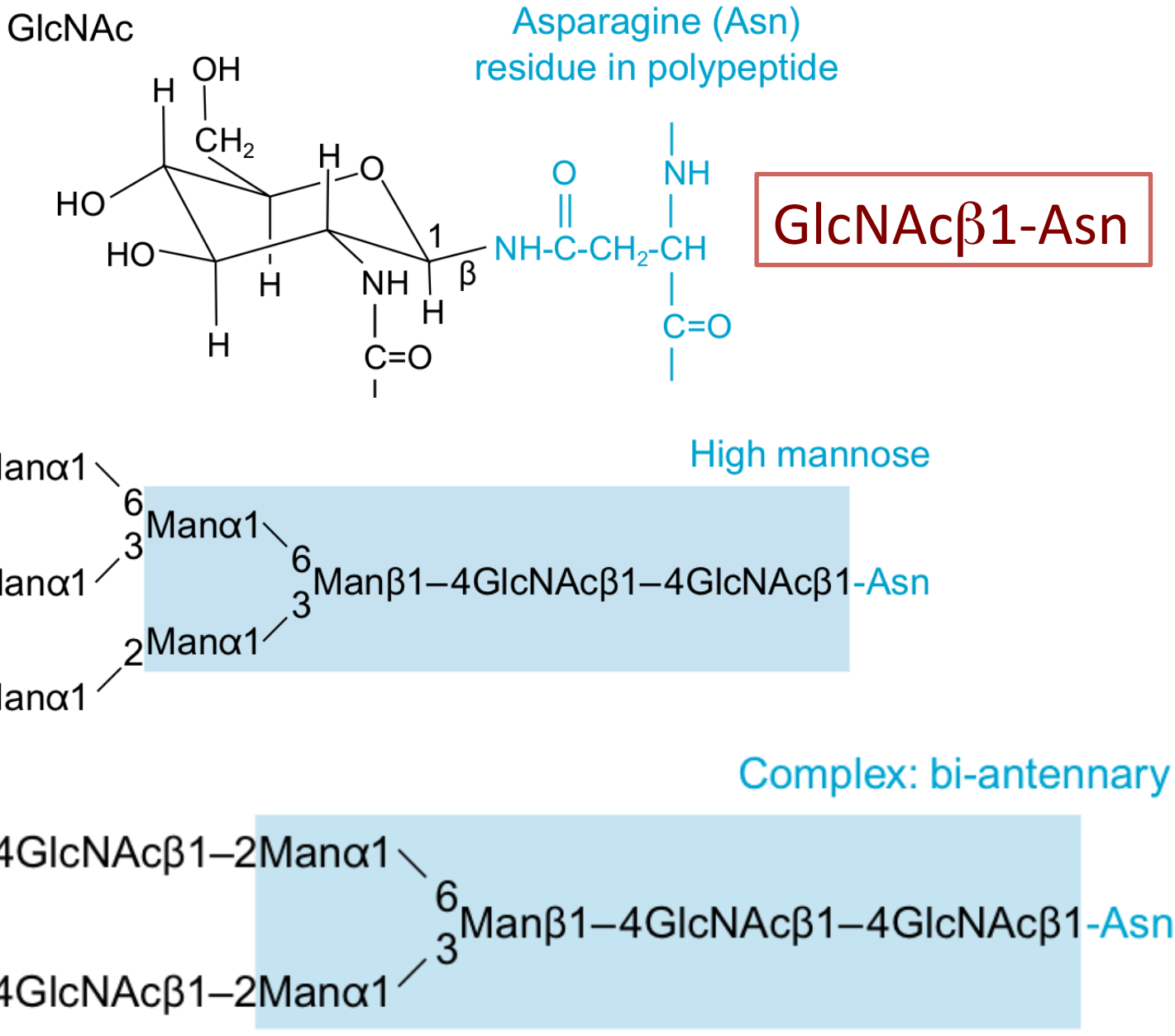
1. N-linked glycosylation

Learning objectives

- The structures of different classes of N-linked glycans
- Chemical steps in the biosynthesis of N-linked glycans
- the cellular compartmentalization biosynthetic steps
- The nature and causes of heterogeneity of N-linked glycans

- The N-linked glycosylation pathway is the best understood route to protein glycosylation.
- This pathway can be used to illustrate general principles of glycoconjugate structure and biosynthesis.
- The nature of the N-linked glycosylation pathway also provides some insight into the functions and evolutionary history

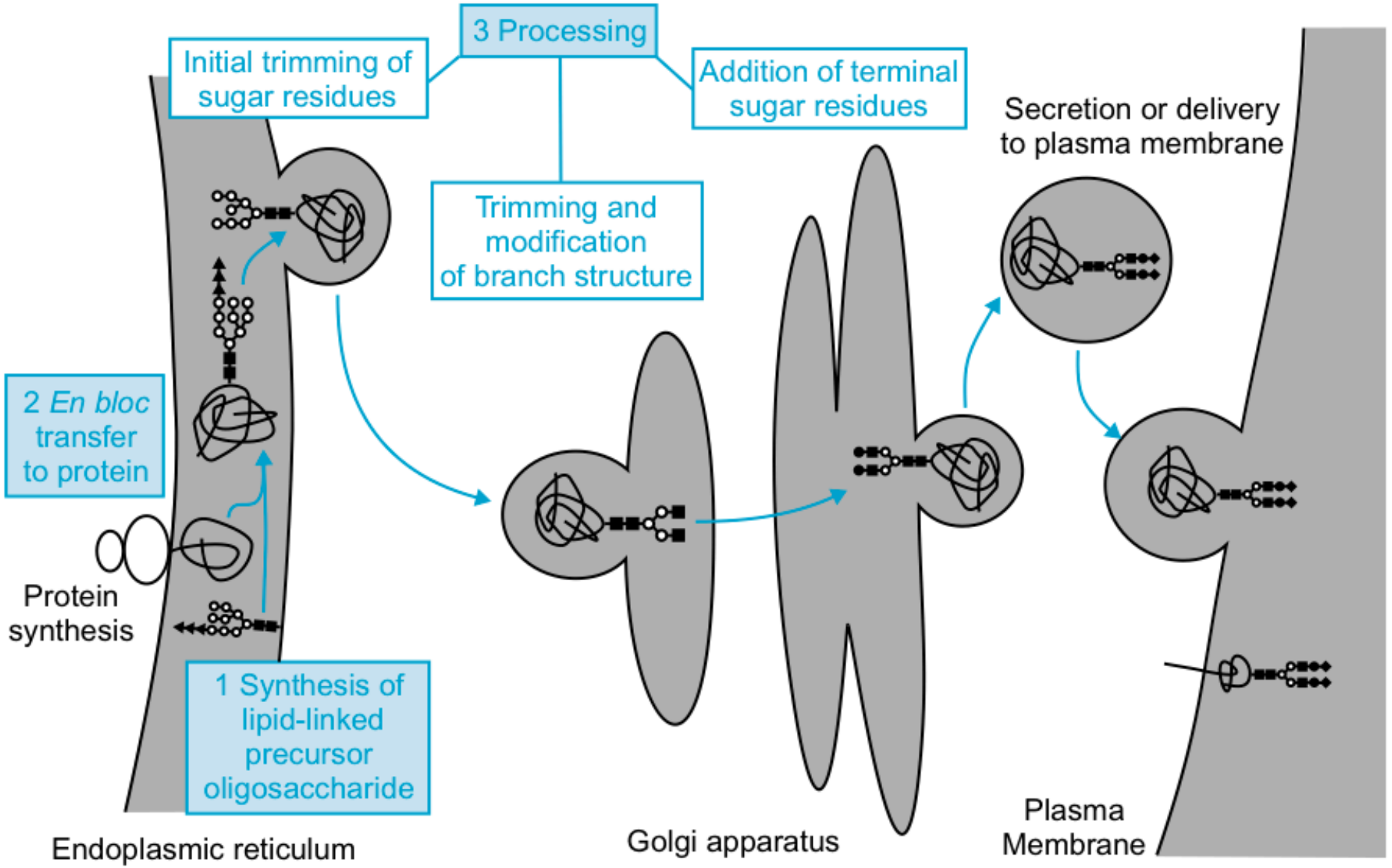
1.1 Diverse N-linked glycans have a common core structure


GlcNAc
Asparagine (Asn)
residue in polypeptide
GlcNAcβ1-Asn
High mannose
Manα1–2Manα1
Manα1–2Manα1
Manα1–2Manα1
Manα1
Manα1
Manβ1–4GlcNAcβ1–4GlcNAcβ1-Asn
Complex: bi-antennary
NeuAcα2–6Galβ1–4GlcNAcβ1–2Manα1
NeuAcα2–6Galβ1–4GlcNAcβ1–2Manα1
Manβ1–4GlcNAcβ1–4GlcNAcβ1-Asn

1.2 Assembly of N-linked glycans occurs in three major stages

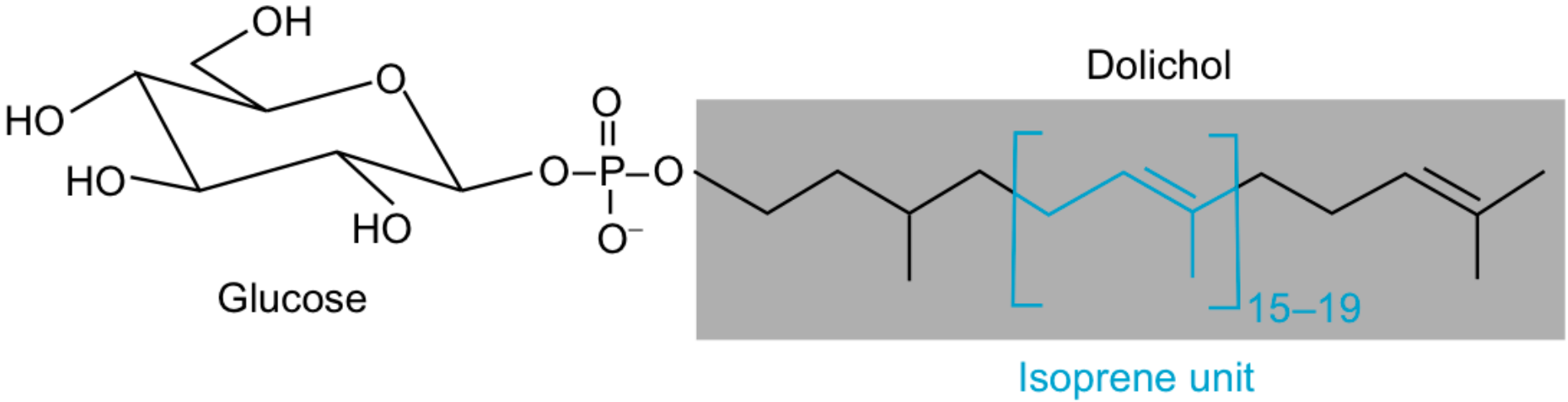
Three stages

- Formation of a lipid-linked precursor oligosaccharide
- *en bloc (总的)* transfer of the oligosaccharide to the polypeptide
- Processing of the oligosaccharide


3 Processing
Initial trimming of sugar residues
Addition of terminal sugar residues
Trimming and modification of branch structure
Secretion or delivery to plasma membrane
2 *En bloc* transfer to protein
Protein synthesis
1 Synthesis of lipid-linked precursor oligosaccharide
Endoplasmic reticulum
Golgi apparatus
Plasma Membrane

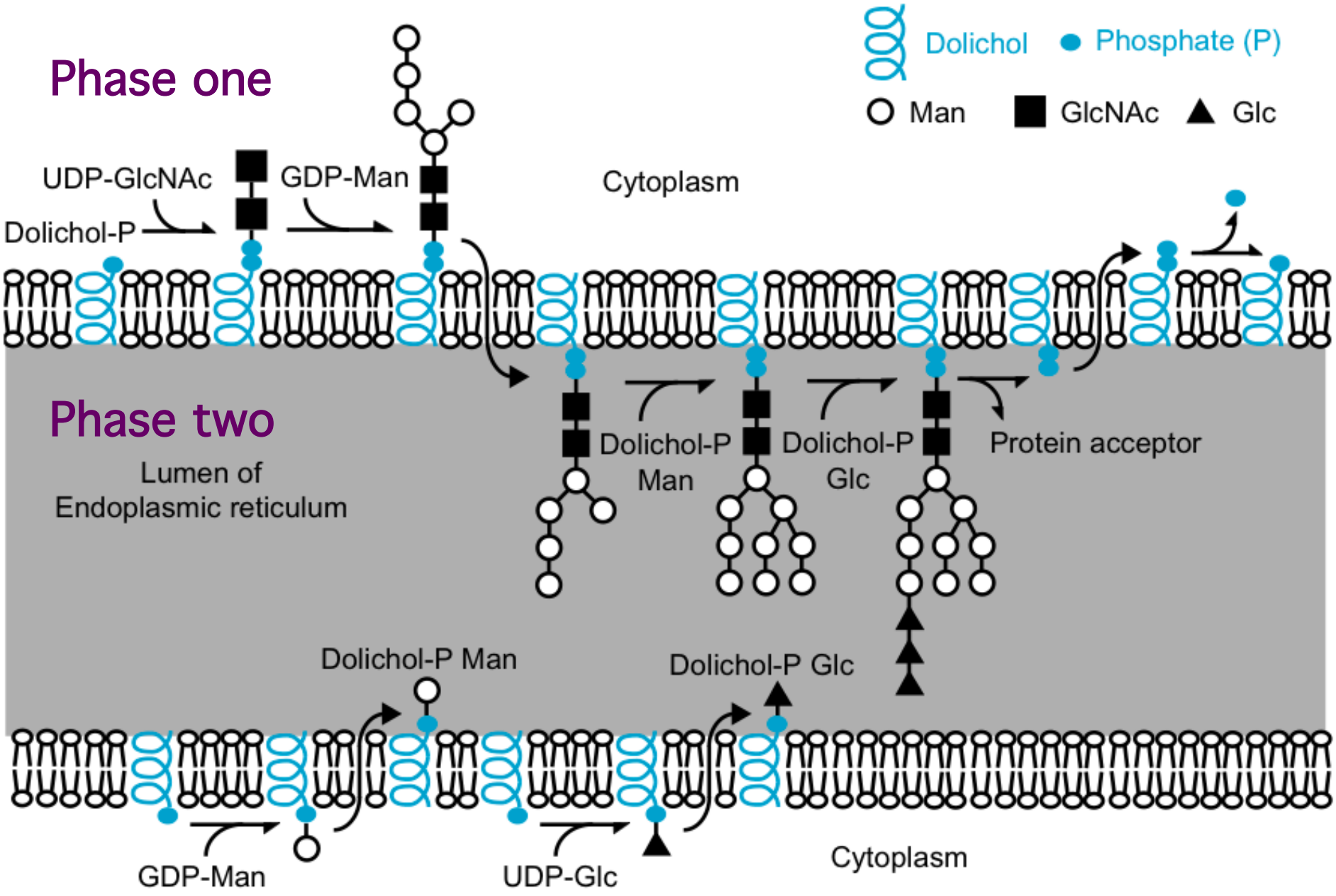
1.3 The precursor oligosaccharide for N-linked glycans is assembled on the lipid dolichol

A dolichol sugar that can serve as the donor in glycosyltransferase rxns on the luminal (腔的) side of the endoplasmic reticulum (内质网) membrane



Hydrophobic portion of this lipid is far longer than the fatty acid tails on membrane phospholipids

Possibly in a helical or folded conformation


Dolichol
Phosphate (P)
Man
GlcNAc
Glc
Phase one
UDP-GlcNAc
GDP-Man
Dolichol-P
Cytoplasm
Phase two
Lumen of
Endoplasmic reticulum
Dolichol-P
Man
Dolichol-P
Glc
Protein acceptor
Dolichol-P Man
Dolichol-P Glc
GDP-Man
UDP-Glc
Cytoplasm

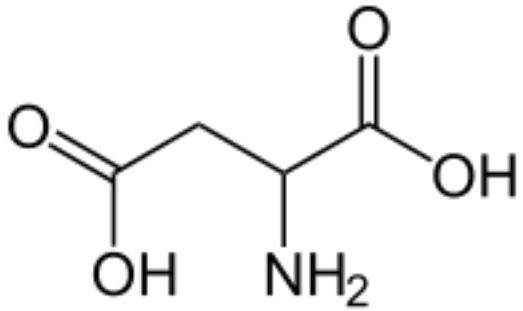
1.4 The dolichol-linked precursor oligosaccharide is transferred to asparagine residues of polypeptides

Three conditions for sugar-attached Asparagine residue

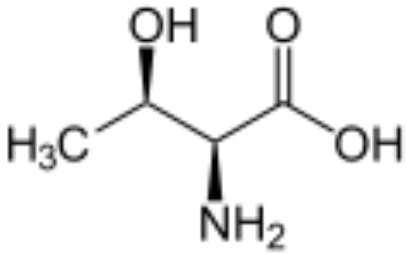
- They must be located in a specific sequence context within the primary structure of the protein
- They must be located appropriately in the 3D structure of the protein
- They must be found in the correct intracellular compartment

Asn-Xaa-Thr or Asn-Xaa-Ser

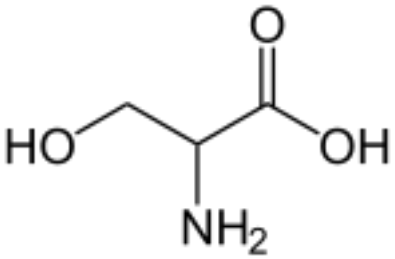
Asn-Xaa-Cys



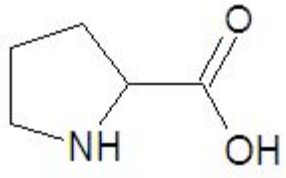
Asn



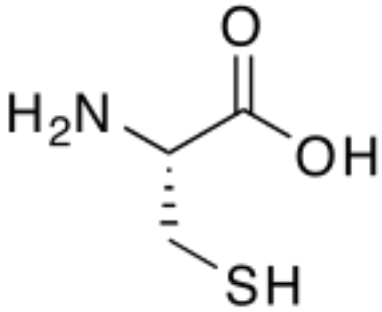
Thr



Ser

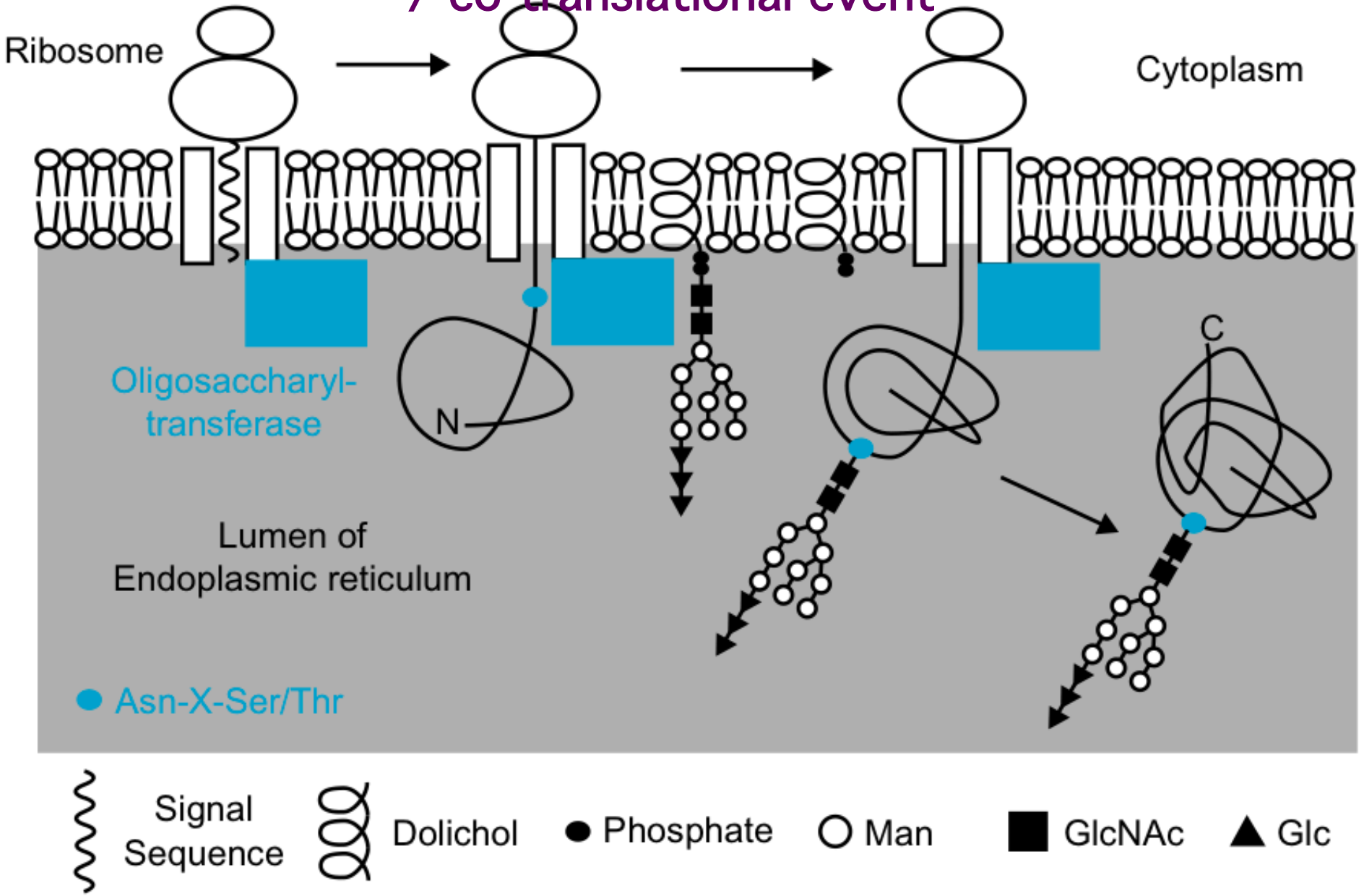


Pro



Cys

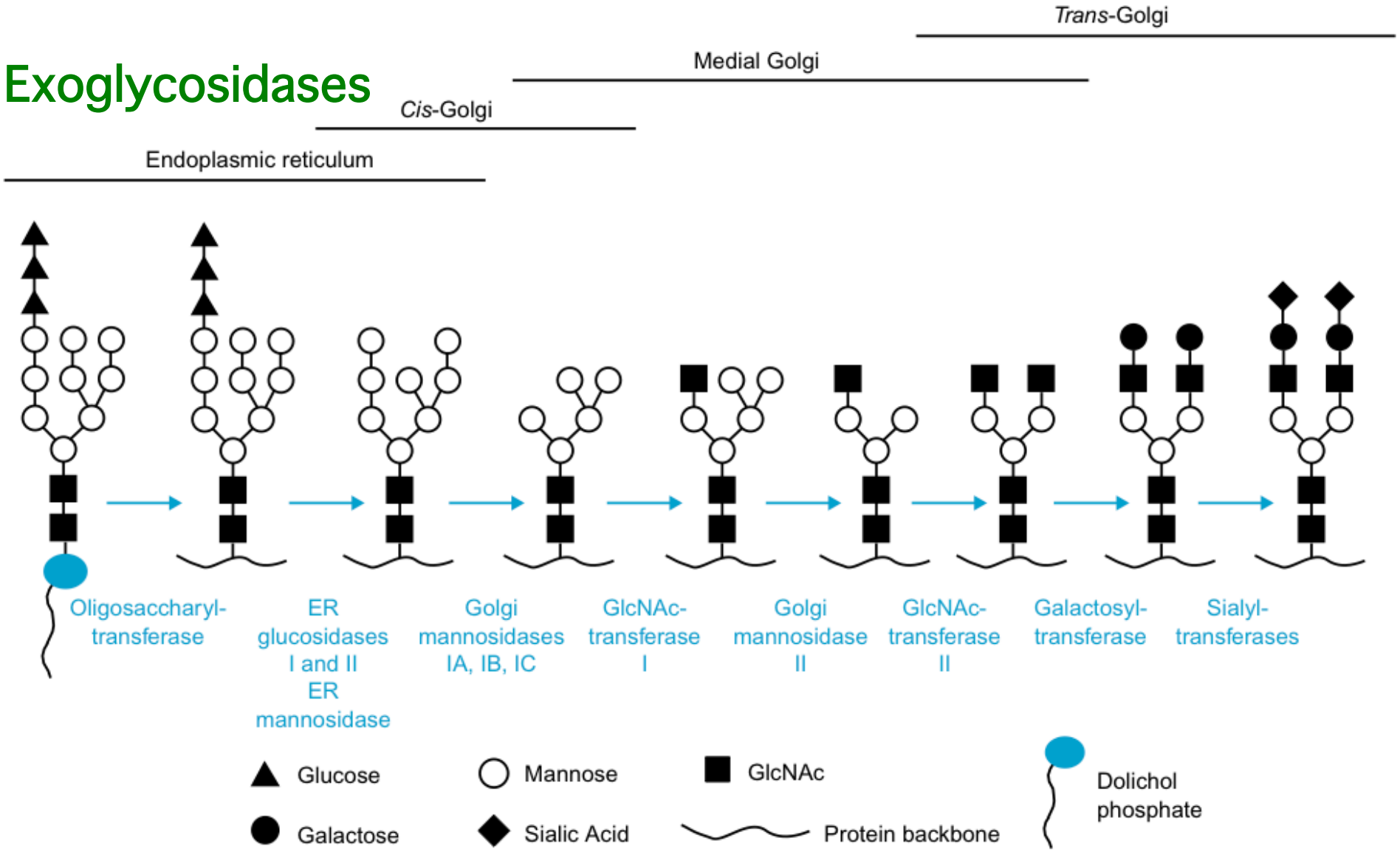
Hydrophilicity / energy sense / protein surface / co-translational event



1.5 The core oligosacchride structure is modified by glcyosidases and glycosyltransferase

Processing of N-linked glycan

Exoglycosidases



Common complex N-linked glycans

High-mannose oligosaccharide

Rare glycans containing five or more branches

Addition of a single Gal and Sialic acid residue to each GlcNAc

Galactosyltransferase sialyltransferase
Locate further along in the secretory pathway
In the trans portion of the Golgi apparatus (高尔基体)
and the trans-Golgi network

Galactose: β1-4 linkage
Terminal sialic acid: α2-3 or α2-6 linkage

Resulting N-acetylneuraminic acid-galactose (NeuAc-Gal)
Capping structure

Donors ------ nucleotide sugars

UDP-GlcNAc, UDP-Gal, cytidine monophosphate (CMP)-NeuAc

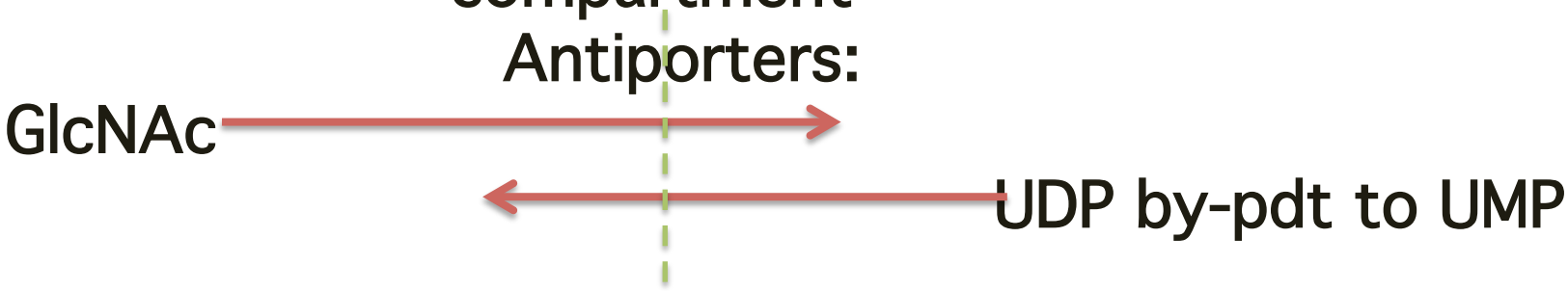
Topological issue

All rxns occur in the lumen of the Golgi apparatus, yet the donors are made in the cytoplasm

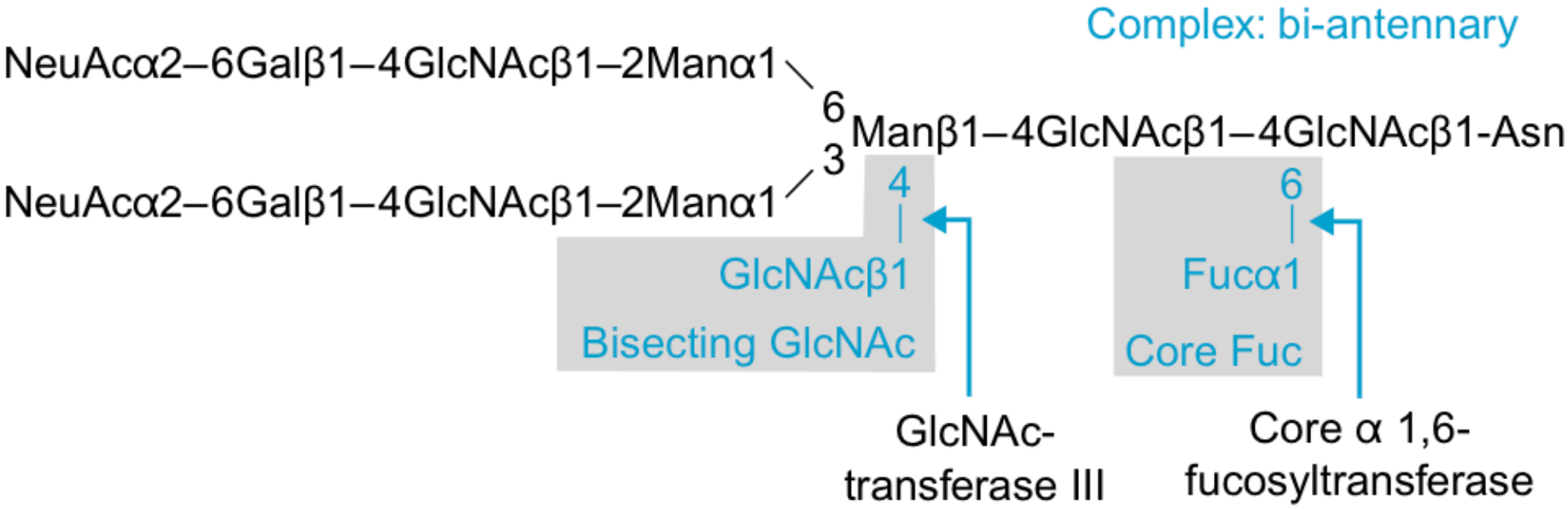
No translocation intermediate analogous to dolichol phosphomannose used in dolichol precursor biosynthesis

Transportors: faciliate entry of the donors into the luminal compartment

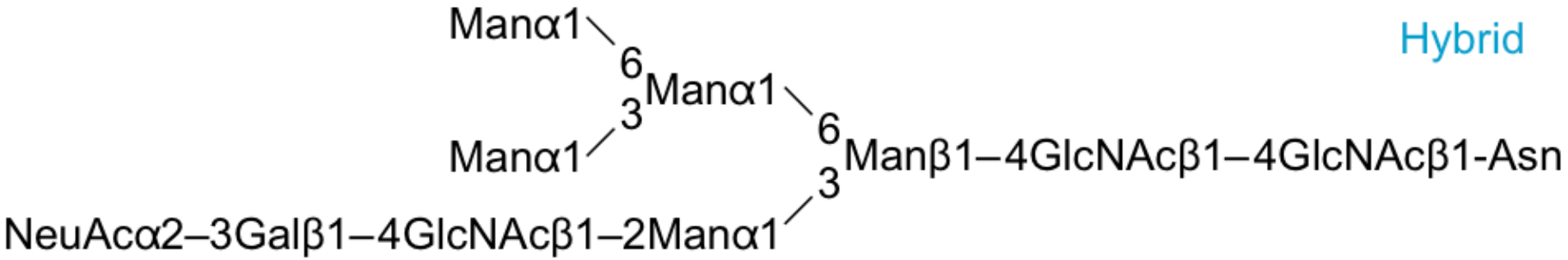
Antiporters:



1.6 Hybrid structures and polylactosamine extensions are common variations of the core oligosaccharide



Golgi processing mannosidases don't act after GlcNAc-transferase I

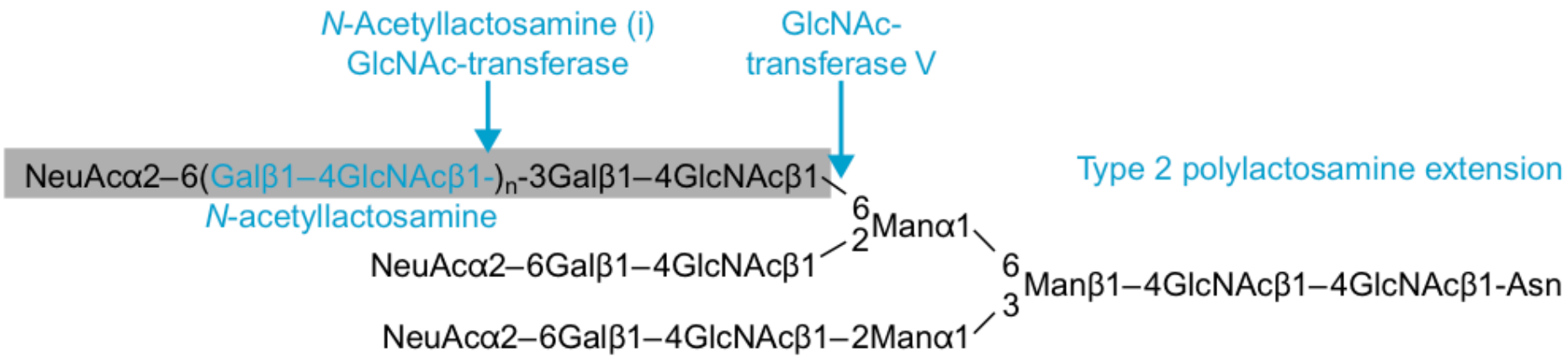


Poly-lactosamine extension

Repeated Gal-GlcNAc disaccharides
Various linkages, most commonly
Galβ1-4GlcNAc and GlcNAcβ1-3Gal

GlcNAc transferase, which can use Gal-terminated branch as an acceptor

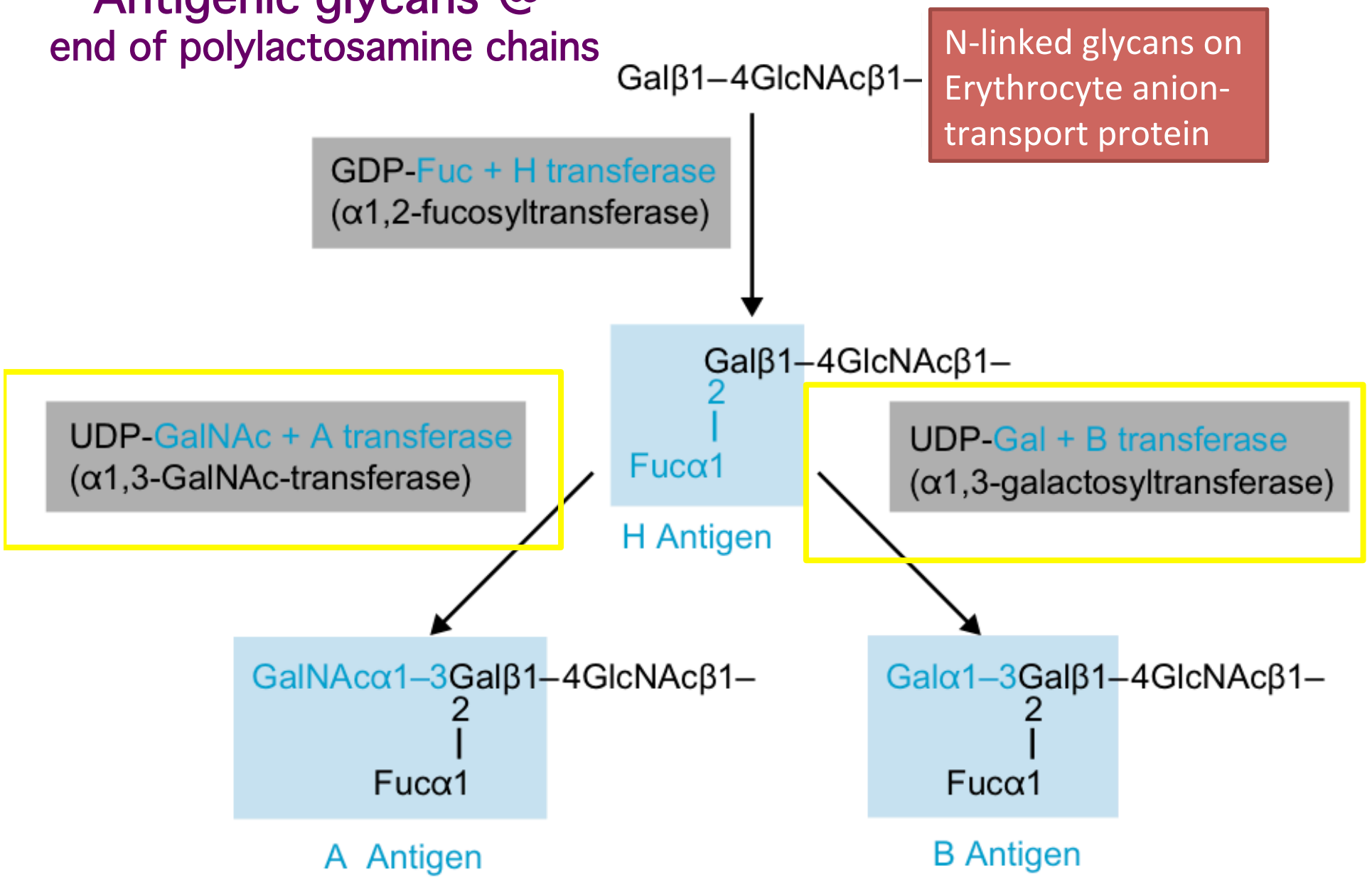
Then repeated action of the GlcNAc-transferase and a galactosyltransferase



Can be Galβ1-3GlcNAc (type 1)

1.7 ABO blood groups are determined by the presence of different terminal sugars on glycans of red blood cells

Antigenic glycans @
end of polylactosamine chains



Glycosyltransferases responsible for addition of Gal and GalNAc

Are encoded by different alleles (等位基因) at a single genetic locus (基因座) on chromosome 9
Galβ1-4GlcNAc and GlcNAcβ1-3Gal

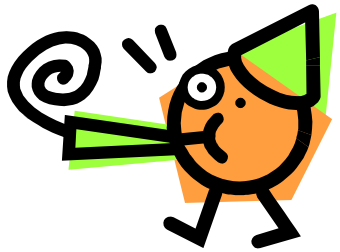
O-type individuals, null alleles are found at this locus in both copies of chromosome 9, so there is no functional transferase and these individuals express only the H substance.

A- type allele encodes a GalNAc-transferase and the B allele encodes a galactosyltransferase

The allelic GalNAc- and Gal-transferases are extremely similar to each other: there is only one essential amino acid difference

Blood Groups And Transfusion

Glycobiology and disease



All individuals are exposed to the A, B, and H structures from an early age, mostly in food substances,

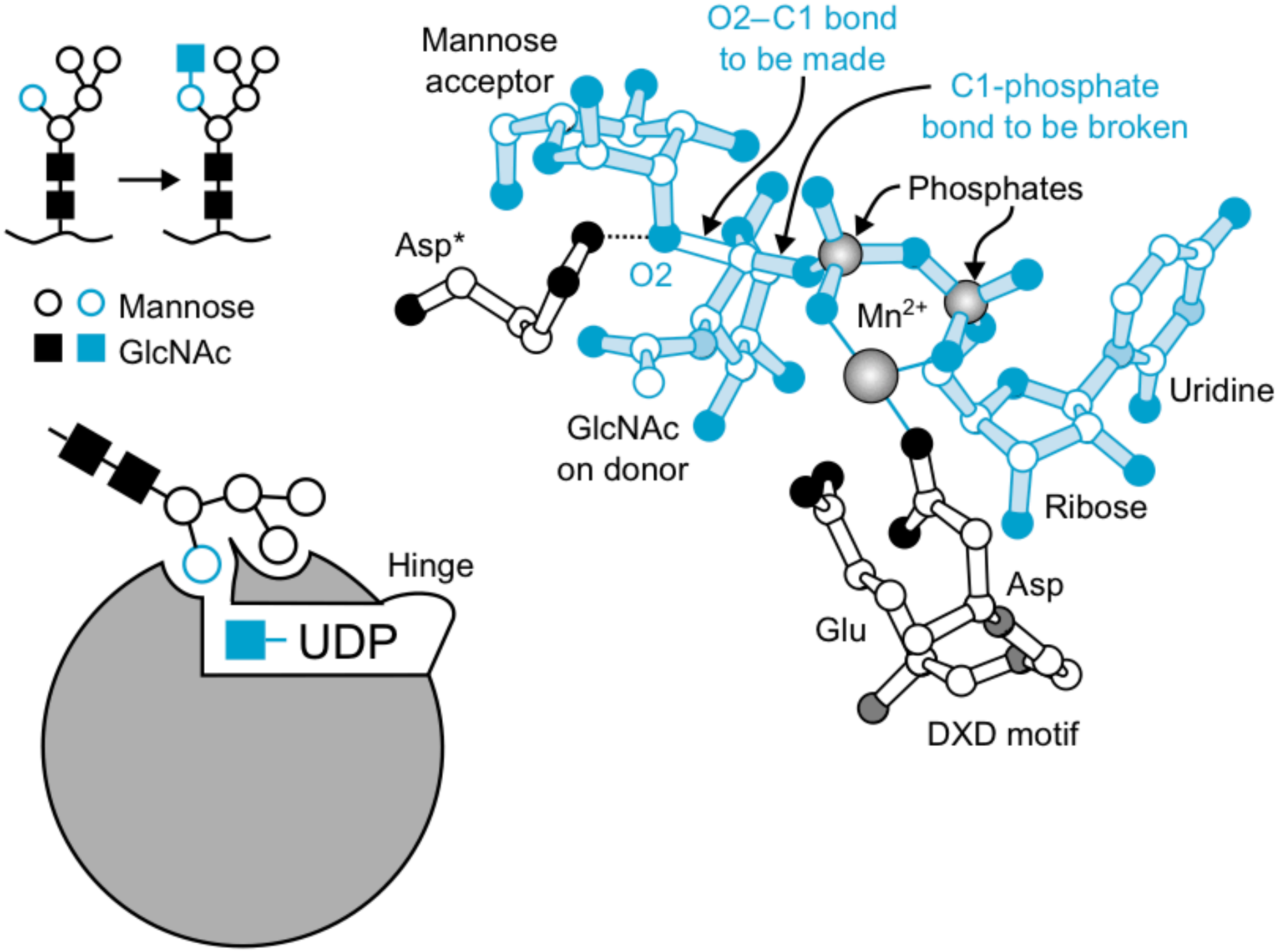
But the immune response is modulated by the nature of the structures expressed on each individual's own cells

A-type individuals develop antibodies against the B structure but are tolerant to the A structure

Transfused cells must not express glycans to which the recipient has antibodies

The H substance was originally defined antigenically because certain rare individuals lack the fucosyltransferase and hence do not make H, A, B structures-------Bombay phenotype (表型)

1.8 Hundreds of glycosyltransferases generate highly diverse N-linked glycans


Mannose
GlcNAc
Hinge
UDP
Mannose acceptor
O2–C1 bond to be made
C1-phosphate bond to be broken
Phosphates
Asp*
O2
Mn2+
GlcNAc on donor
Uridine
Ribose
Asp
Glu
DXD motif

1.9 The N-linked glycans of an individual glycoprotein are usually heterogeneous

Lack of correspondence between glycan and protein structures

Asp residue in different copies of a glycoprotein can be modified with different glycans

Glycoprotein molecules with a common polypeptide chain but bearing different glycans are called glycoforms

>1000 possible glycans

Several thousand possible glycoforms

Typically fewer than 100 glycoforms of most glycoproteins and usually only a few of these predominate

1.10 The nature of N-linked glycans attached to an individual glycoprotein is determined by the protein and the cell in which it is expressed

- The structure of a protein can have some influence on its glycosylation
- Common N-linked structures are found on many glycoproteins as defaults
- But certain terminal elaborations are found only on subsets of glycoproteins
- Modifications of glycans partly specified by structural features of the protein
- E.g. Man-6-phosphate, GalNAc 4-SO_4, blood group

- Position of certain glycosylation sites favours processing of N-linked glycans to complex structures
- The same protein expressed in different cell types, or in a single cell type under different growth conditions, can be glycosylated in different ways
- May rise from different glycosyltransferase
- Or reflect the different rates at which glycoproteins transit through the secretory pathway
- How long they encounter various secretory pathway

Tumor cells express different glycosylation compared to normal cells derived from the same tissues

1.11 High-mannose structure are present in lower eukaryotes, but the glycosylation machinery has evolved to produce complex glycans in higher organisms

- The route for synthesis of the complex glycans seems particularly circuitous
- The core structure of mammalian glycoproteins built of mannose and GlcNAc are often the end pdts in simpler eukaryotes
- In yeast and fungi, the core structures are extended rather than processed and they serve a structure role, forming the outer wall of these single-celled eukaryotes
- High-Man structure represents an evolutionary precursor to the more complex terminal elaborations needed for recognition

- An initially uniform core is present on the glycoproteins as they transit the early part of the secretory pathway and subsequent modification of the core approaching cell surface
- **The glycans can serve different functions inside the cell, related to protein folding and intracellular trafficking, and outside the cell (involving recognition events)**

1.12 N-linked glycans are essential for the development of multicellular organisms

Phenotypes of some key knockout mice

- Mutations in the dolichol pathway that prevent synthesis of the core oligosaccharide can be used to generate cells that lack N-linked glycosylation
- Such cells can be made, N-linked glycans perform no function that is indispensable for cell viability
- But they are not capable of forming an organism
- Embryos lacking N-linked glycans perish shortly after implantation

- High-mannose oligosaccharides are not enough to allow normal development, because embryos lacking GlcNAc-transferase I, develop only to embryonic day 10
- Importance of N-linked glycosylation in multicellular organisms

2. O-linked glycosylation

Learning objectives

- The structure of mucin-type glycans and the ways that they are organized on polypeptides
- The structures and biosynthesis of glycosaminoglycans and their attachment to core polypeptides in proteoglycans
- The functions of mucins and proteoglycans in water retention and tissue structure
- The unusual features of cytoplasmic and nuclear glycosylation

- The properties of two groups of glycoprotein, mucins and proteoglycans are dominated by the large number of O-linked sugars
- Physical properties appropriate for the important biological functions that they perform
- Additional types that have signalling functions

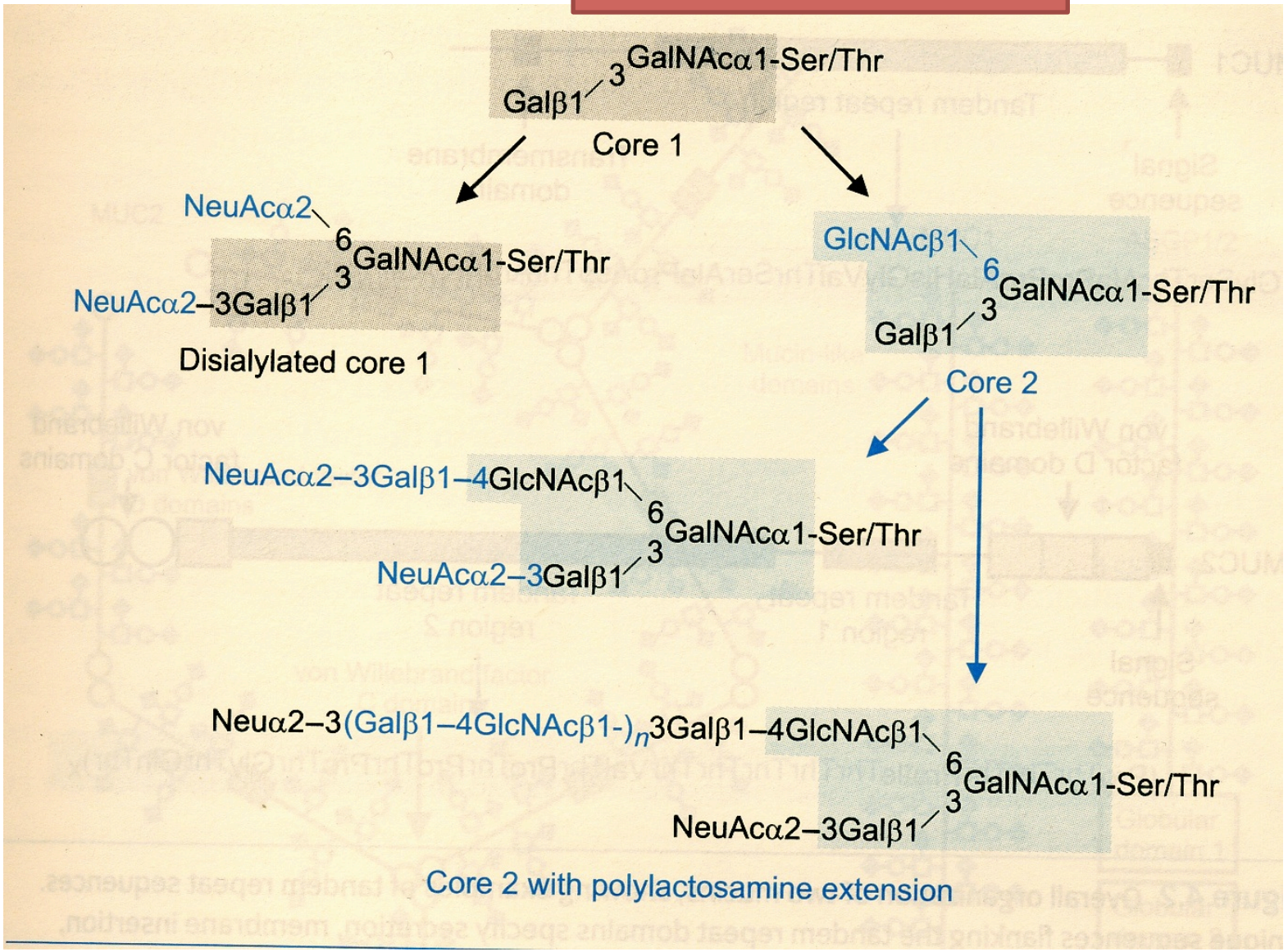
2.1 Mucins are large, heavily O-glycosylated proteins that hold water

The primary purpose: to retain water

- at surfaces that are exposed to the environment but are not sealed by moisture-impermeable layers, as in the skin
- Present at the surfaces of the digestive and genital tracts and in the respiratory system
- Salivary, intestinal, vaginal, nasal mucins
- Characterized by different polypeptide backbone

- Addition of clusters of sialylated glycans to groups of serine and threonine residues results in regions of strong negative charge
- Give mucins the capacity to bind large amounts of water
- Highly hydrated, energetic cost of releasing water from mucins reduced the rate of evaporation

GalNAcα1-Ser/Thr

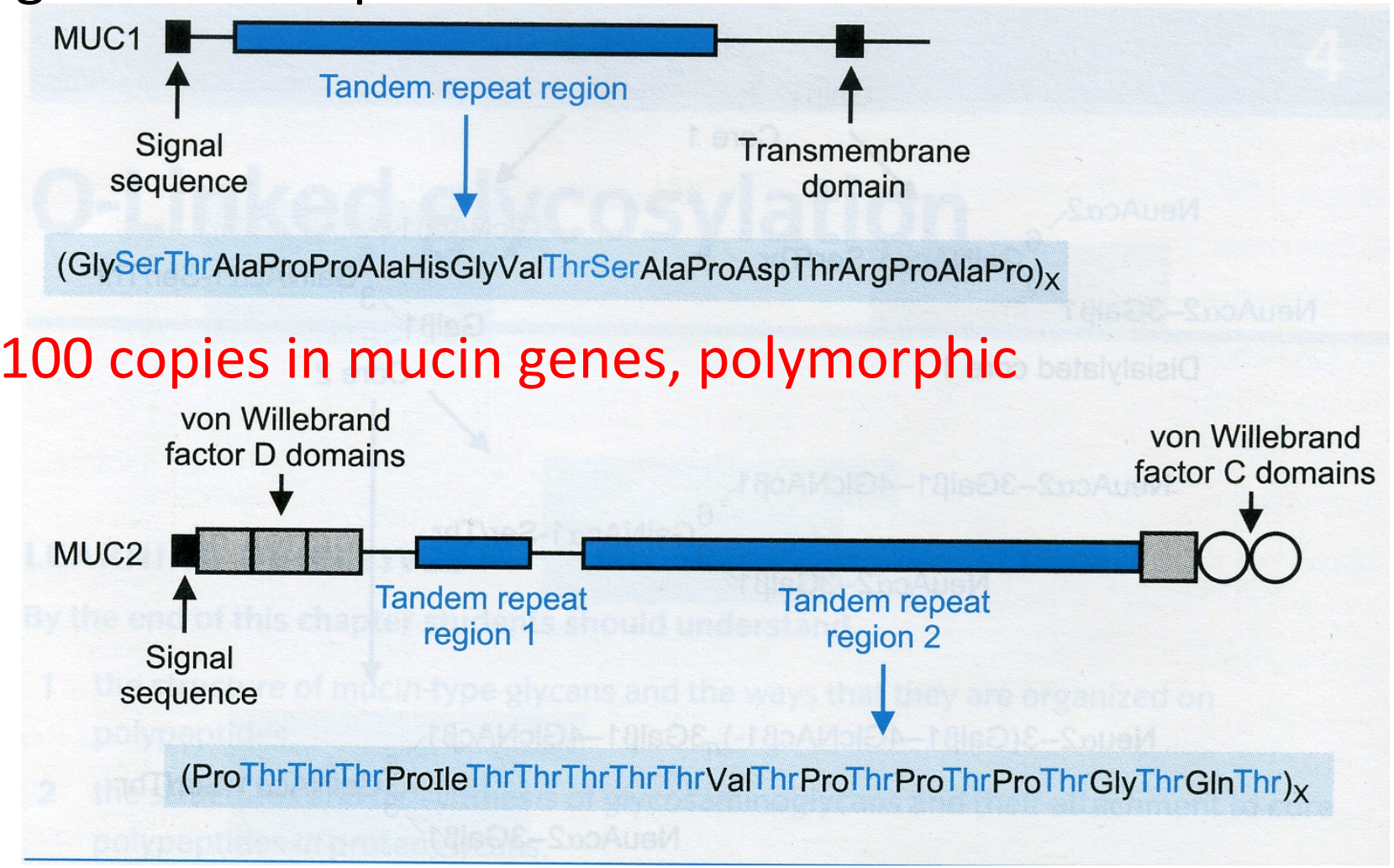


At least eight genes encoding core polypeptides in humans

- In addition to aiding in water retention
- The mucins serve as lubrication and also help to protect from invasion by microorganisms
- The organization of a mucin polypeptide reflects these various functions
- Polypeptides are exceptionally long, containing as many as 10K AA
- Can be membrane bound or secreted

Tandem repeats of relatively simple AA sequences rich in Ser and Thr

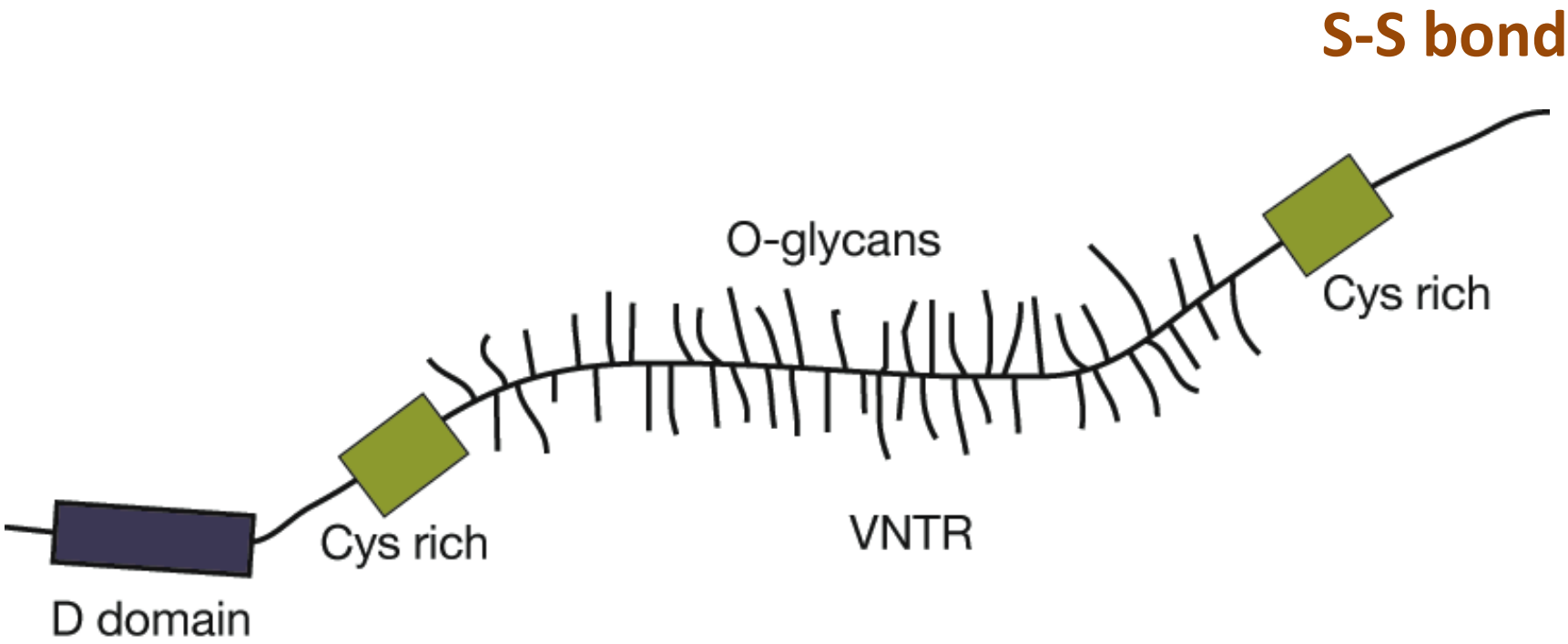
Identical or nearly identical within a specific mucin gene in one species



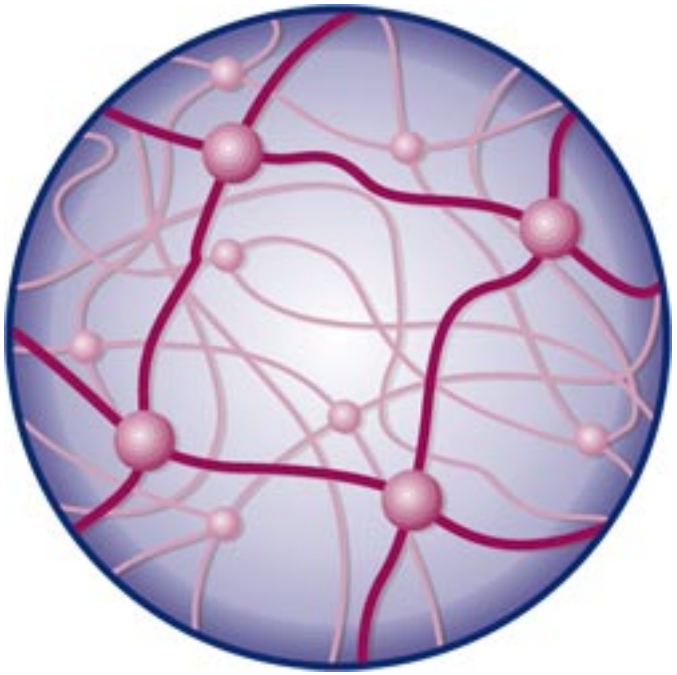
> 100 copies in mucin genes, polymorphic

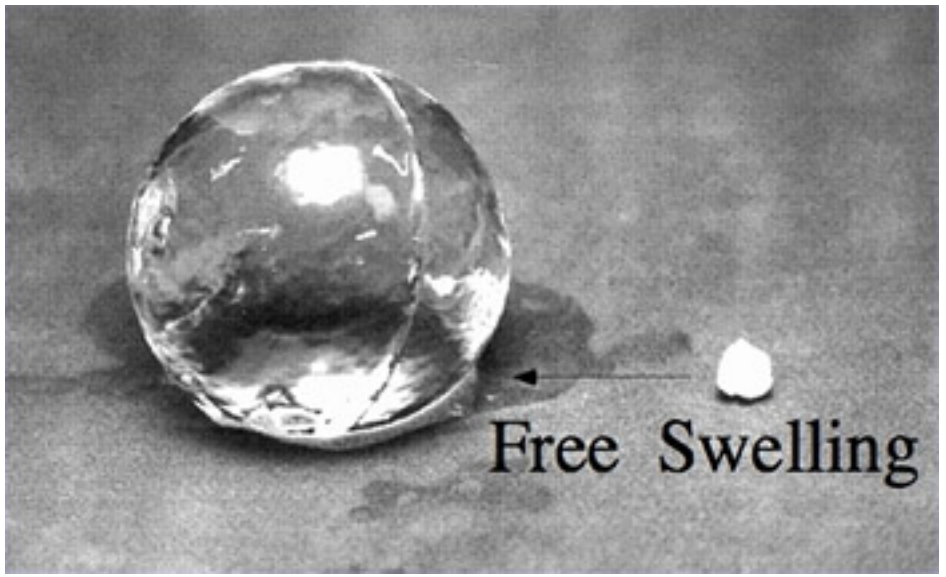
- The overall sizes of the mucin molecules are increased by formation of disulphide-linked oligomers
- Globular terminal domains in secreted mucin polypeptides direct formation of covalent complexes larger than 1MDa
- Dense glycosylation of the mucins leads to an extended polypeptide conformation
- Accommodate the bulk glycans and charge repulsion
- The large size and elongated dimensions of mucin oligomers gives the hydrated protein viscoelastic properties
- Essential for lubrication functions , form gels

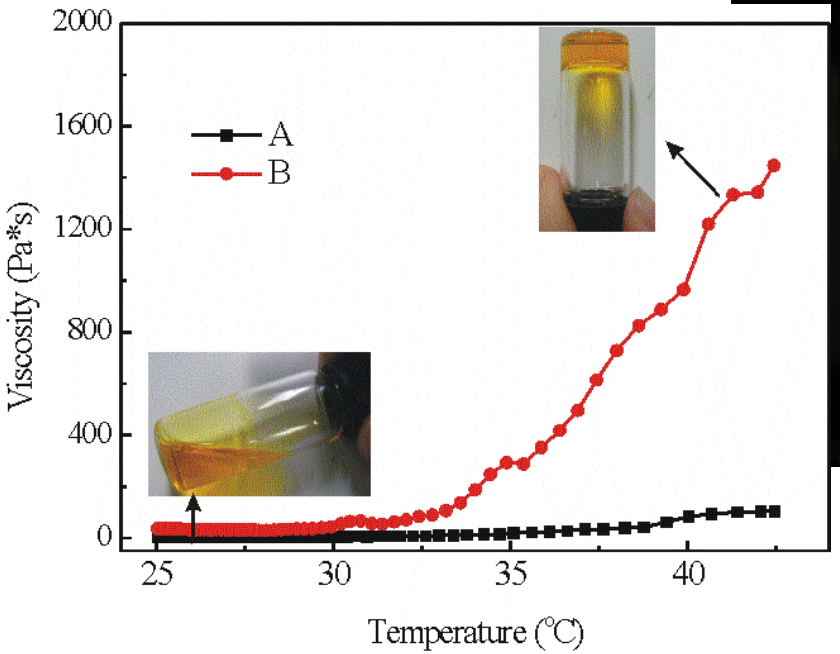
A simplified model of a large secreted mucin

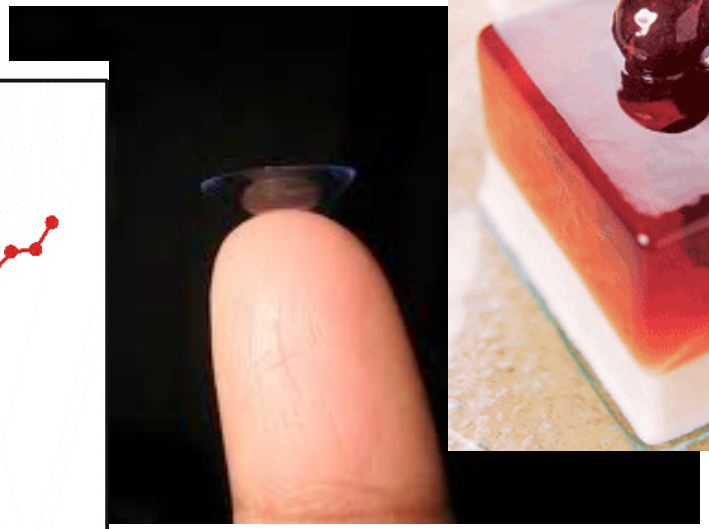


variable number of tandem repeat









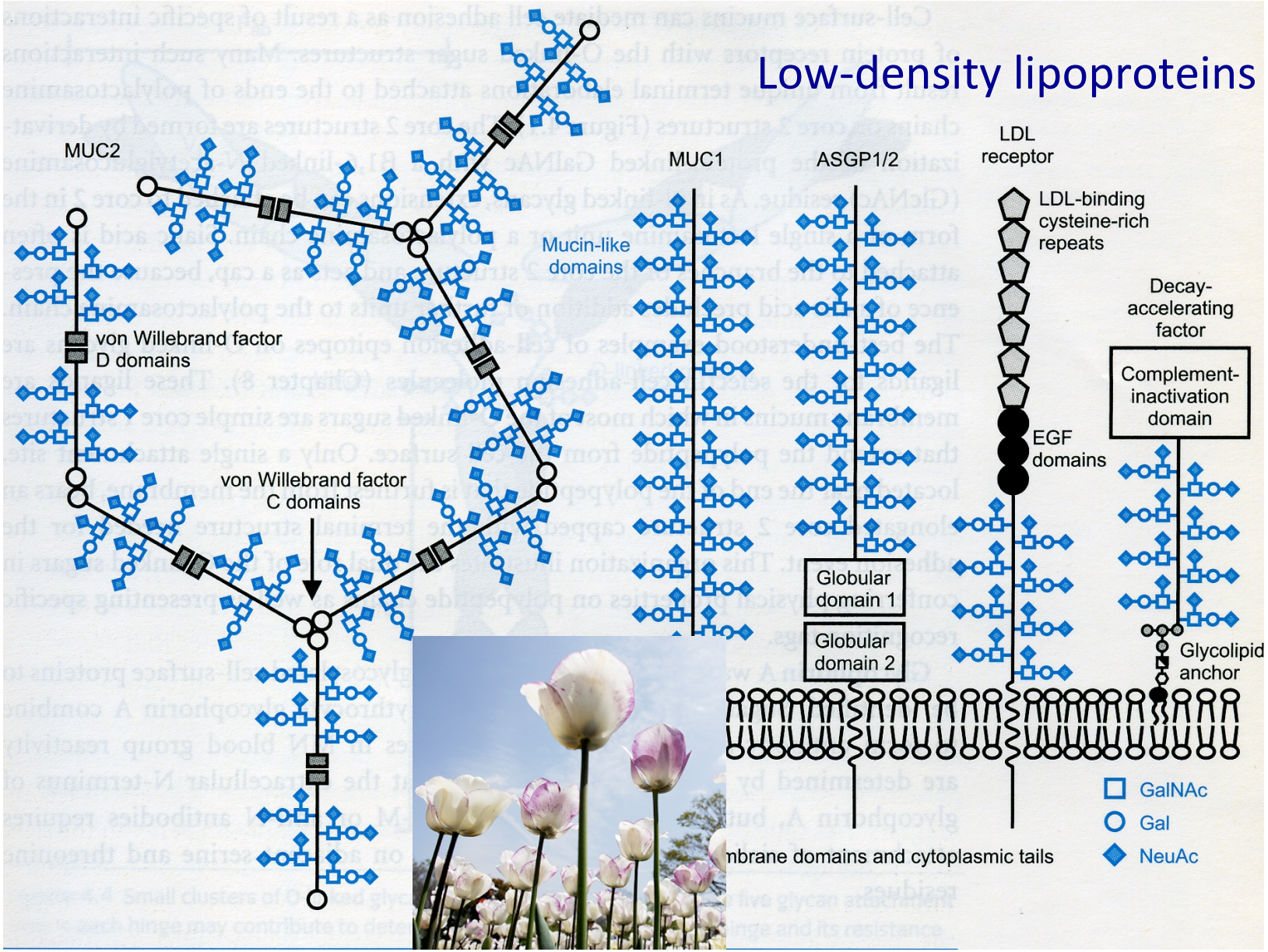
pilokolok.t4h.info

Antibacterial properties

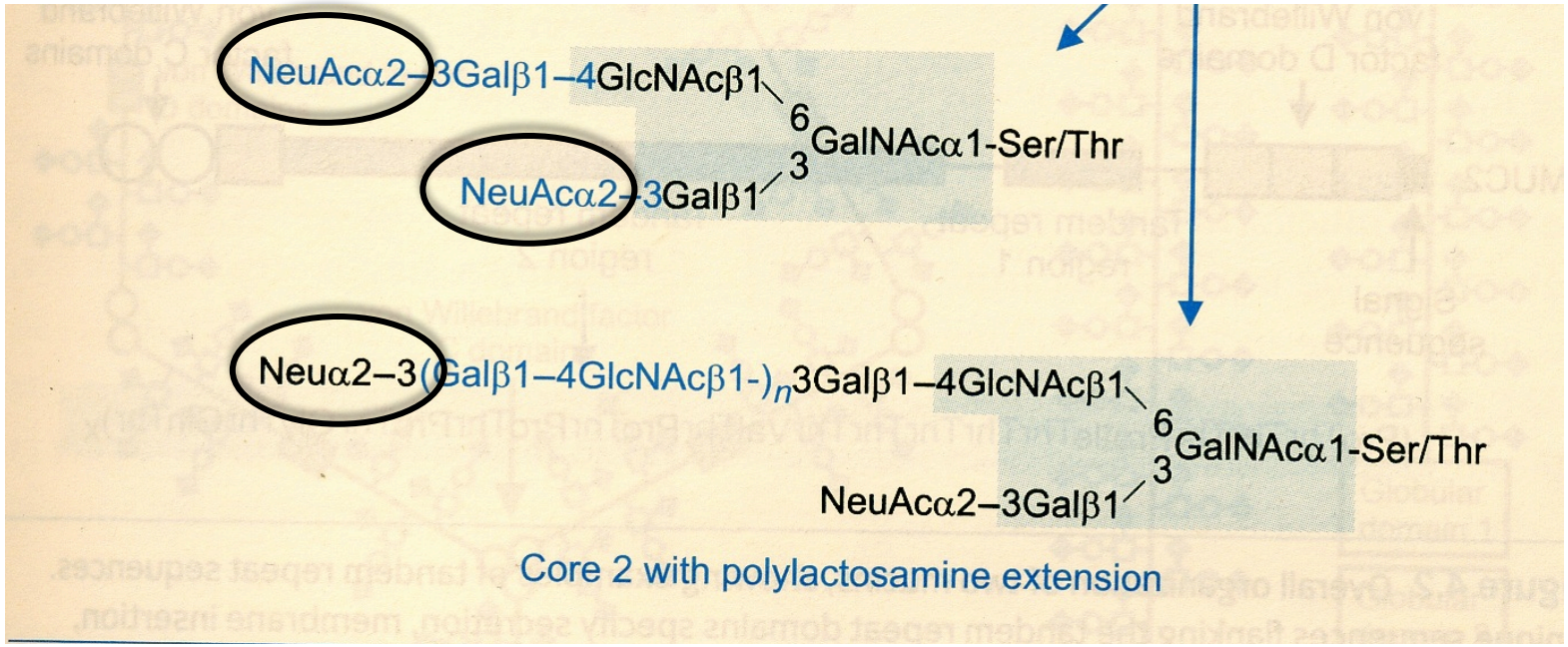
- The mechanical properties of viscous mucin solutions and gels can serve to entrap potential pathogens and clear them
- Some terminal domains have a specific affinity for bacterial cell surfaces, which can enhance entrapment and may also lead to immune responses against the bound organisms
- Cell-surface mucins can regulate adhesion between cells

2.2 Some cell-surface proteins have mucin-like domains

Low-density lipoproteins



Mediate cell-adhesion

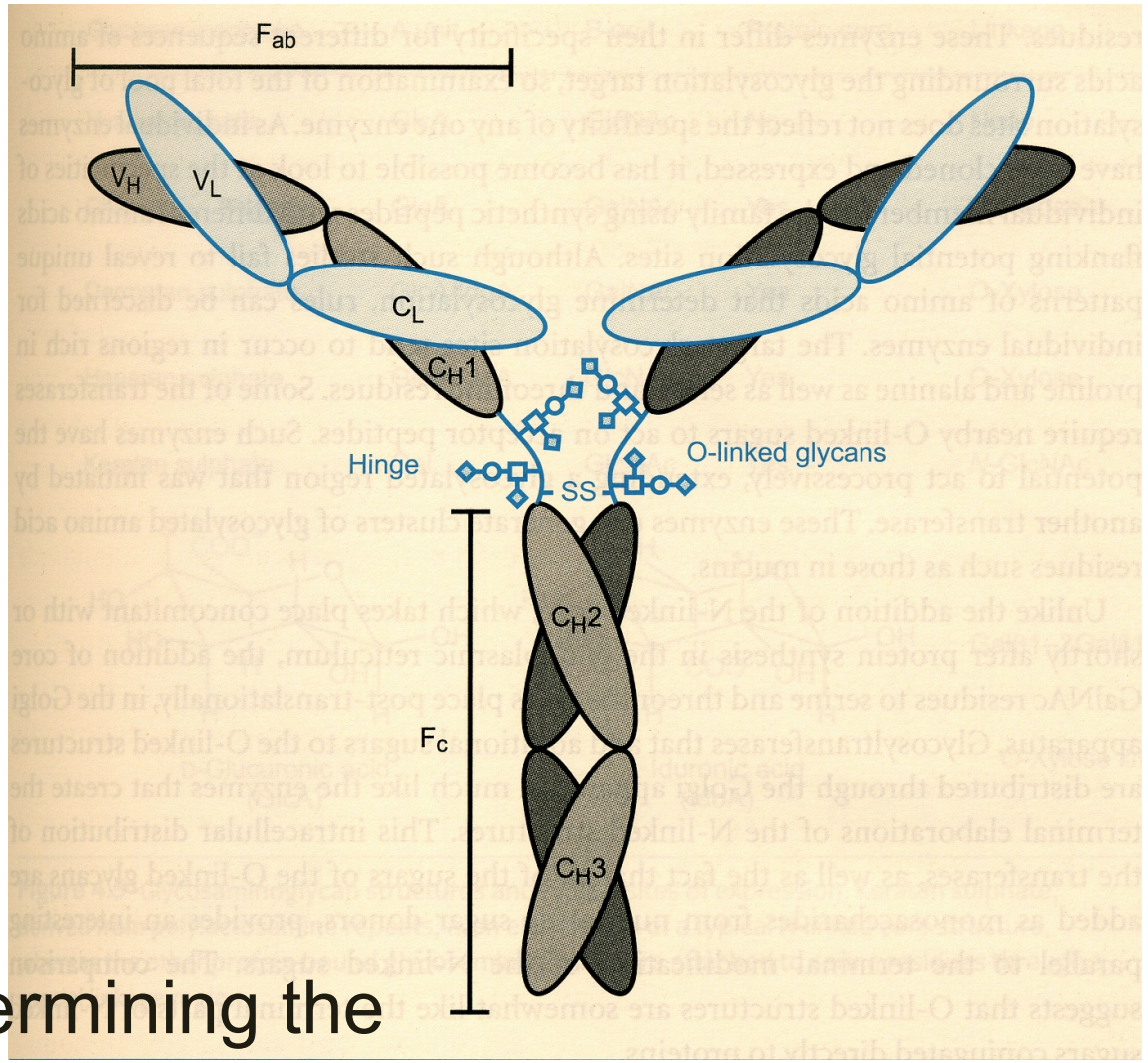


e.g. Cell-adhesion epitopes on O-linked glycans are ligands for the selectin cell-adhesion molecules

Membrane mucins, most of the O-linked are core 1 structures, with a elongated core 2 structure attached

2.3 Many soluble and cell-surface glycoproteins contain small clusters of O-linked sugars

Small clusters of O-linked glycans in the hinge region of IgA



contribute to determining the conformation and resistance to proteolysis

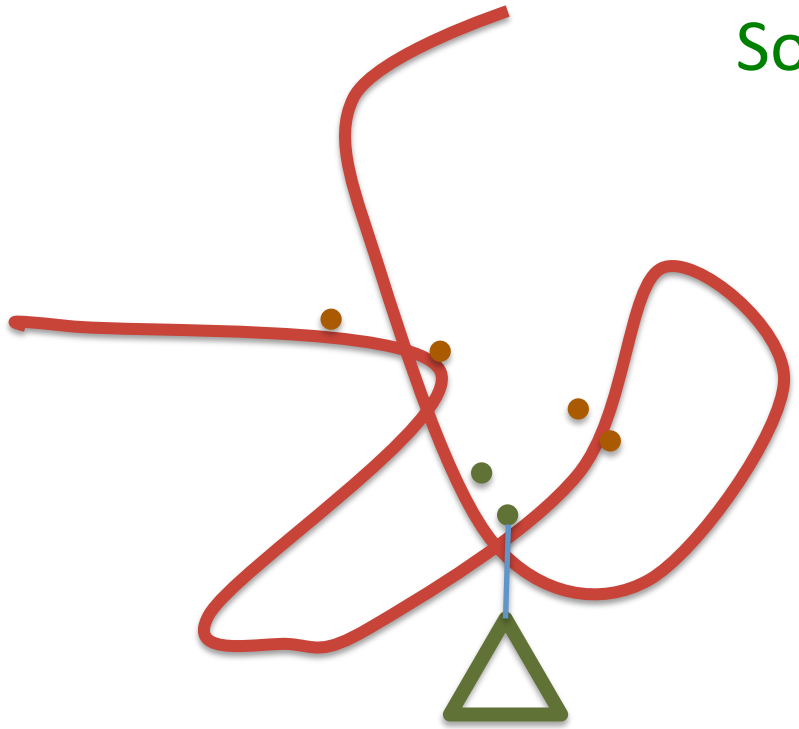
2.4 Biosynthesis of mucin-type sugars occurs by sequential addition of monosaccharides to proteins in the Golgi apparatus

O-glycosylation machinery utilizes analogous glycosyltransferases for N-linked with different organization

- All sugars are added one at a time in a stepwise series of rxns, starting with the first GalNAc
- No performed core or en bloc transfer
- No simple target sequences for O-linked glycosylation analogous to the Asn-XaaSer/Thr sequences

- Multiple rxns for the lack of a consensus O-glycosylation sequence
- Numerous transferases can attach GalNAc to serine and threonine
- Enzymes differ in their specificity for different sequences of AA surrounding the glycosylation target
- No reflection of the specificity
- Rules can be discerned for individual enzymes

Regions rich in Proline and alanine as well as serine and threonine residues



Some requires nearby O-linked sugars to act on acceptor peptides

Act processively, extending a glycosylated region that was initiated by another transferase, generate clusters of glycosylated residues

- Addition of core GalNAc residues to serine and threonine takes place post-translationally
- In the Golgi apparatus
- Other glycosyltransferases are distributed through the Golgi apparatus
- Like the enzymes that ceate the terminal elaborations of the N-linked structures
- O-linked structures are somewhat like the terminal parts of N-linked sugars conjugated directly to proteins