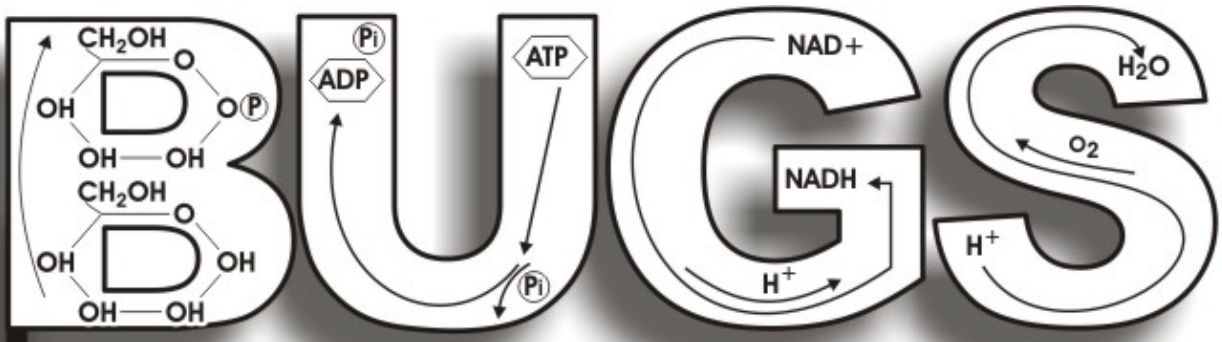
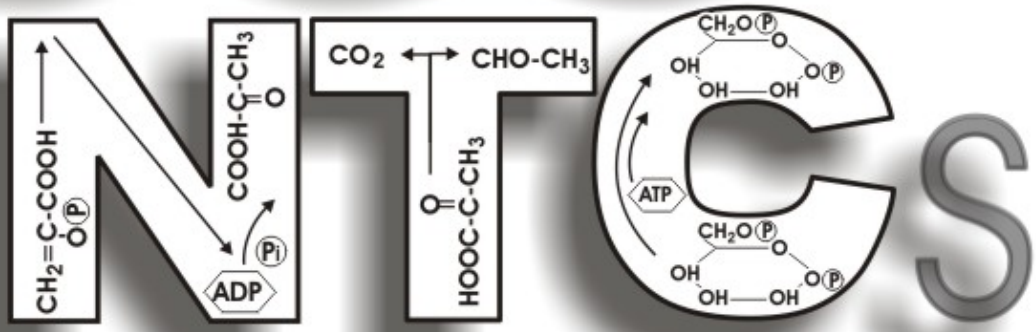




BIOC 312

SET #10 Lectures 32- 33 (04/02/07 –04/04/07)

ANNOUNCEMENTS

< Biochemistry Undergraduate Society - 5th Floor McIntyre rm. 511 >
< bugs@sus.mcgill.ca - (514) 398-5247>

04/02/2007

Lecture #32– Protein Glycosylation I

NOTE: This NTC is meant to be used as a study aid to supplement your own class notes. Hence, not all of the text or diagrams contained in the lecture slides will be reproduced here.

Please send any comments or questions about NTCs to either Misha or James through e-mail at **bugs_u2reps@yahoo.ca**

Introduction:

- If you have any questions you can email Dr. Annette Herscovics at annette.herscovics@mcgill.ca
- The textbook can be used as a reference but not everything covered in the lecture is found in the textbook.
- We are encouraged to look through a biochemistry textbook and read the chapter on sugars.

Overview of Glycosylation:

Protein glycosylation is one of the major post-translationally modification. This area of biochemistry is part of a field know as **Glycobiology**, which covers many macromolecules, but we will only cover glycoproteins, and we will not be covering proteoglycans which are a special kind of glycoproteins.

- There are two basic types of glycosylation of proteins:

1. Glycosylation of membrane and secreted glycoproteins → these are made within the secretory pathway.
2. Cytosolic and nuclear glycosylation.

- During this modification there is a **covalent** linkage of the carbohydrate to specific amino acids, the most common ones are **asparagines, serine and threonine**.
- This covalent linkage is present in **glycol-peptide linkages** and the most common linkage is ***N*-linked oligosaccharides** (also known as *N*-glycans).
- This linkage is the N-acetlyglucosamine attached to asparagines.
- There are also *O*-glycans, since the carbohydrate is linked with either serine or threonine.
- Glycoproteins may have variable amounts of carbohydrates (little to a lot); sometimes it can make up 90% of the molecular mass.
- Carbohydrates are quite hydrophilic; therefore, carbohydrate moieties on glycoproteins would largely be located on the outside of the protein and not in the hydrophobic core.
- Because the synthesis of the carbohydrates on a protein is not a template-mediated process, it entirely depends on the specificity of the enzymes that are involved in the synthesis and the availability of substrates.
- Instead there is an assembly line type of mechanism, whereby each enzyme works in sequence, dependent on the previous enzyme's work. This allows more room for variation.
- As a result of this we have a concept known as **microheterogeneity**, which means that at each glycosylation site you might have a given carbohydrate structure with different degrees of completion of that carbohydrate structure; at each site you can have a mixture of oligosaccharides.
- As a result glycoproteins exist as **glycoforms**, note that there is no such thing as a pure glycoprotein, because there are variations present in the carbohydrate structure. This is important because when the glycoprotein in a cell has a certain biological function, variations in the carbohydrate structure can influence this.
- The functions of glycoproteins vary and it depends on the type of glycoprotein.
- Some of the functions of glycoproteins are entirely structural and in some cases the carbohydrate can play a role in the conformation of the protein and its ability to fold. In other cases the carbohydrate moiety can effect a proteins solubility (the more carbohydrate there is on a protein, the more soluble the

glycoprotein will be). In some cases if there is a lot of carbohydrate some physical properties, such as viscosity, will be affected.

- Besides a structural role there are very specific functions of carbohydrates as recognition molecules. By their nature (variety) there are many recognition functions that they can have depending on the structure of the carbohydrate.
- These functions of recognition are largely mediated by proteins that are called **lectins** or by glycoproteins themselves. Their property is to be very specific for certain carbohydrate structures. There is a whole family of lectins which recognize different carbohydrate structures and thereby modulate all kinds of interactions and confer biological properties.

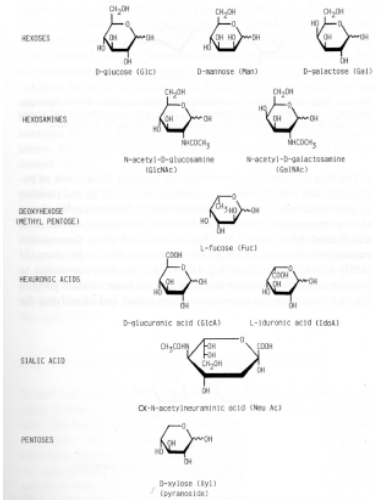
Cytosolic/Nuclear O-GlcNac:

- *O*-GlcNac is an abbreviation for the *O*-*N*-acetylglucosamine.
- In the cytosol or nucleus a modification occurs where a single residue of *N*-acetylglucosamine is attached to serine or threonine.
- In this type of modification the sugar is never substituted for another sugar. Only a single sugar residue, which occurs on many cytosolic and nuclear proteins.
  - *Example:* transcription factors, RNA polymerase II, cytoskeletal proteins, oncogenes, nuclear pore proteins (particularly rich in O-GlcNac).
- One of the characteristics of this modification (O-GlcNac0 is that it is very **dynamic**; there is an enzyme that adds N-acetylglucosamine from UDP-GlcNac to serine or threonine residues and there is an enzyme that readily removes the N-acetylglucosamine.
- There is possibility of a high turnover of N-acetylglucosamine. There is a great deal of research in this field and one postulation is that it may be a modulator of phosphorylation.
- A researcher at Johns Hopkins postulated that the O-GlcNac modification undergoes a mechanism known as the **ying-yang model**, which states that because the GlcNac is located at or near sites of phosphorylation (either directly on the serine or threonine that is normally phosphorylated or indirectly). The ying-yang model says that the O-GlcNac is playing a role in phosphorylation because there is a competition between glycosylation and phosphorylation of that specific residue. The residue will have one or the other depending on the functions.
- These proteins have a mostly regulatory function.

Glycosylation Within the Secretory Pathway:

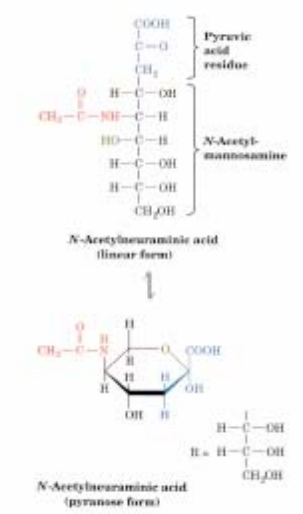
- These glycoproteins within the secretory pathway (Golgi apparatus or the Endoplasmic Reticulum) can be soluble glycoproteins or membrane glycoproteins. The peptide moiety is made on membrane-bound ribosomes, due to the presence of a signal sequence.
- The addition of the carbohydrate to these proteins occurs entirely on the **luminal** side of the ER and the Golgi, because the addition of the carbohydrate molecule is post-translational, as the peptide moiety is being made on the membrane, simultaneously the addition of the carbohydrate can occur.
- The types of glycoproteins are very varied. Some are soluble and secreted (they end up in the extracellular milieu), yet others are targeted to intracellular organelles such as lysosomes and Golgi. Importantly, there are membrane glycoproteins that are targeted to the plasma membrane. These exist as part of the plasma membrane.
- On the plasma membrane, generally, the sugar molecules face the outside of the cell. This occurs because glycoproteins synthesis occurs on the luminal side, and this is the side of the ER or Golgi which is incorporated into the plasma membrane, in this specific orientation.
- The carbohydrate extends to the extracellular environment of the cells, and it is important to note that all cells have a carbohydrate coat. If you stain for carbohydrates (by light microscopy or electron microscopy) you will see a specific coat, which is carbohydrate-rich.

- The carbohydrate molecule is very important at the surface of cells because it is capable, by its recognition functions, to mediate the interaction of cells with their environment. Cells interact with:
  - Other cells
  - Intracellular adhesion
  - Extracellular matrix
  - Soluble molecules present in the serum



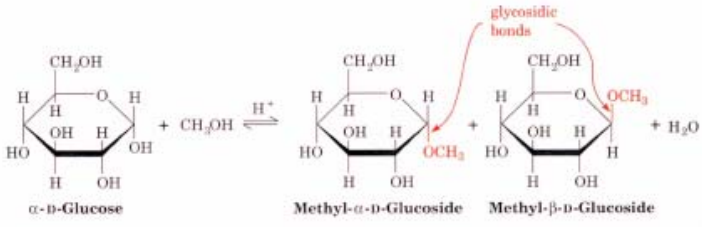
Sugar Structures:

- The most important sugars present on glycoprotein molecules in animal cells are seen on page 3 of the handout.
- The structures that we need to know: **glucose**, **mannose** (differs from glucose in the 2-position), and **galactose** (differs from glucose in the 4-position).
- **Hexoses** are represented in the howarth configuration (not a 3D configuration) in a **pyranose form** (6-membered ring).
- **Hexosamines** are derivatives of glucose and galactose; the amino group is present on the **2-position**; note that it is usually acetylated.
- **L-fucose** is a pentose and shown in the D,L-configuration, this is normally a terminal molecule on a carbohydrate chain.
- **Hexuronic acid** are the sugars which have been oxidized at the 6-position have a carboxylic acid.
  - Example: glucoronic acid or iduronic acid
- **Sialic acids**, the major component in humans of this family is α-*N*-acetylneuraminic acid (9-carbon sugar). This basic structure allows for all sorts of substituent groups and can therefore form a family.
- **Xylose** is present on serine or threonine residues in proteoglycans.
- Sialic acid is different from the other 5 or 6 membered sugars. This is really condensation of N-acetyl-mannosamine (6 carbons) with phosphoenolpyruvate (3 carbons of PEP). The way that sialic acid is made the reducing function is on the 2-position in contrast to the other sugars where the reduction function is commonly on the 1-position.



Glycosidic Bonds:

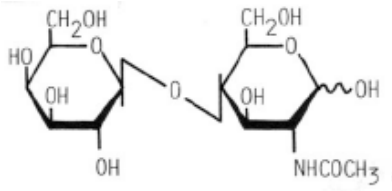
When the cell makes carbohydrate structures and oligosaccharides there is a formation of **glycosidic bonds**. A glycoside is formed by a sugar that has a reducing end; glucose has its reducing end at C1. Elimination of water will occur with any alcohol; in this case it is methanol.
Once the glycosidic bond is formed there is two possibilities (an α or a β isomer) from the two orientations of the molecule during the formation of the glycosidic bond.



It is possible to form a glycosidic bond with every hydroxyl on a sugar giving rise to an α or a β isomer; therefore a lot of variation is possible when two sugars are interacting. This is quite different from proteins, in which a peptide bond between two amino acids can only be formed in one manner. In addition to which the addition of sugar can occur on more than one sugar and therefore linear chains can be formed or branched chains can be formed.

Nomenclature:

The nomenclature is quite difficult, because there are different ways of describing the carbohydrate structures. The structure to the right represents a **disaccharide**, which contains N-acetlyglucosamine linked β-1,5 to galactose. The galactose is linked on the 1-position and the β-galactoside is linked on the 4-position. In biochemical literature the trivial names or condensed name tend to be used the most. The convention states that the reducing end of the sugar is always to the right.



Biosynthesis of Glycoproteins:

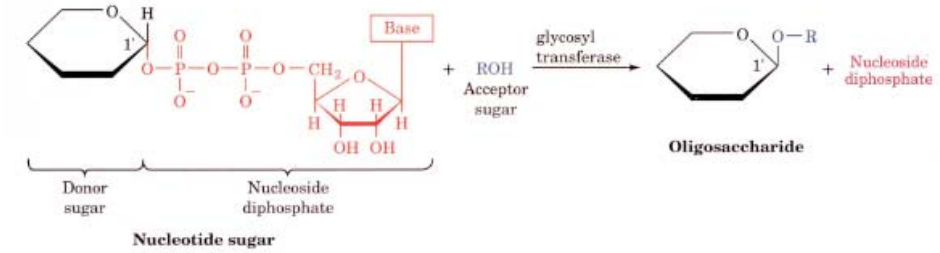
Activated sugar donors

Sugar	Activated form
Glucose, Galactose, N-Acetylglucosamine, N-Acetylgalactosamine, Glucuronic acid, Xylose	UDP-sugar
Mannose, Fucose	GDP-sugar
Sialic acid	CMP-Sia

For biosynthesis of glycoproteins and other carbohydrate structures the sugar moiety must first be activated, free sugars cannot directly act in biosynthetic pathways. They are usually activated through intermediate metabolism, starting from glucose through the glycotlytic pathways to form various nucleotide sugars. The activated forms of the sugars used in biosynthesis are seen in the table seen to the left. The reactions that use these sugars have preferences for the nucleotides.

Glycosyltransferases:

The enzymes that carry out these synthesis reactions are called **glycosytransferases**. The most common type, depicted in the following diagram (not the only type), is the nucleotide sugar which has a sugar donor and the nucleoside diphosphate attached to it. The acceptor can be another sugar or the peptide to which the first sugar will bind to form a **glycol-peptide bond**. The **transferase** will transfer the carbohydrate moiety to the acceptor, to form the new glycosidic bond. Then the nucleoside diphosphate is *released*. In this case we have the transfer of the monosaccharide to the acceptor and the activated sugar is a diphosphate.

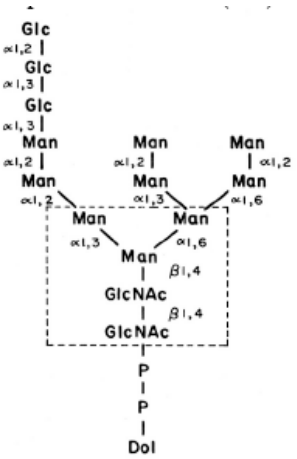


Note that sialic acid transfer occurs from CMP-sialic acid (monophosphate), the same reaction would occur except a nucleoside monophosphate would be released.

Structure of N-Linked Oligosaccharides:

- This is one type of glycoprotein containing **N-glycans**. N-glycans are formed by the addition of carbohydrate through a glycopeptide bond through aspargine and N-acetlyglucosamine.
- They are formed on a very specific sequence which is: **Asp-X-Ser/Thr** (X = any amino acid *except proline*).
- If you look at the sequence of amino acids in a given protein, it is possible to know if this protein will be glycosylated by looking for this sequence, which represents a **potential** glycosylation site for the addition of N-glycan.
- This tripeptide sequence is absolutely essential for N-glycan synthesis, but this does not mean that it is always glycosylated. The glycosylation event depends on where this sequence is located in the given protein.
- An invariant core for N-glycans is a 2-N-acetylglucosamine and mannose and two branches of another mannose. This is seen in the diagram on page 6 of the course hanouts.
- Attached to this invariant core are more mannose residues in the '**high mannose**' structure. In total there are 9 mannose residues; an example of microheterogeneity would be at a given site on a protein you would have a high mannose structure made up of 5,6,7,8 or 9 mannose.

- The **complex** form has branches instead of mannose residues attached to the core branches. The branches do vary to some extent.
- ***Side Note:*** some branches on complex oligosaccharides can be elongated by sugars and become quite large and it has been shown that having increased branching and large oligosaccharides is associated with increased metastatic (allow tumor cells to metastasize) properties of tumor cells.



- Between the Complex and the High Mannose carbohydrates there are **Hybrid strucutures** which retain one branch from the complex and still retains some of the mannose structures.
- The synthesis of N-glycans is a very complicated process with many different steps/enzymes.
- It begins in the endoplasmic reticulum with the synthesis of a precursor molecule, shown to the right. The precursor has the same core as the other N-glycans, and the boxed portion looks like a high-mannose oligosaccharide. The difference between this precursor and the high-mannose form are the three glucose residues attached specifically to the left branch.
- This precursor is formed on a lipid called **dolichol** (Dol). The dolichol-linked oligosaccharides (LLO – lipid linked oligosaccharides) are pyrophosphate-linked carbohydrates.

Dolichol-PP-Oligosaccharide Synthesis

- To make this precursor is a very elaborate biosynthetic scheme that occurs in the endoplasmic reticulum.
- Dolichol is first transformed to dolichol-P by a very specific **CTP-dependent kinase** in the plasma membrane; the phosphate sticks out into the cytosolic side.
- Dolicol-P is important to participate in glycoprotein synthesis; there are two reactions that occur with this molecule.
  - The first reaction involves the transfer of UDP-N-acetylglucosamine of **N-acetyl-glucosamine phosphate.**
  - The second reaction involves the transfer of just the sugar (N-acetyl glucosamine).
- This involves continuous addition of sugars from GDP-mannose and these steps occur in a specific, sequenced manner.

04/04/2007

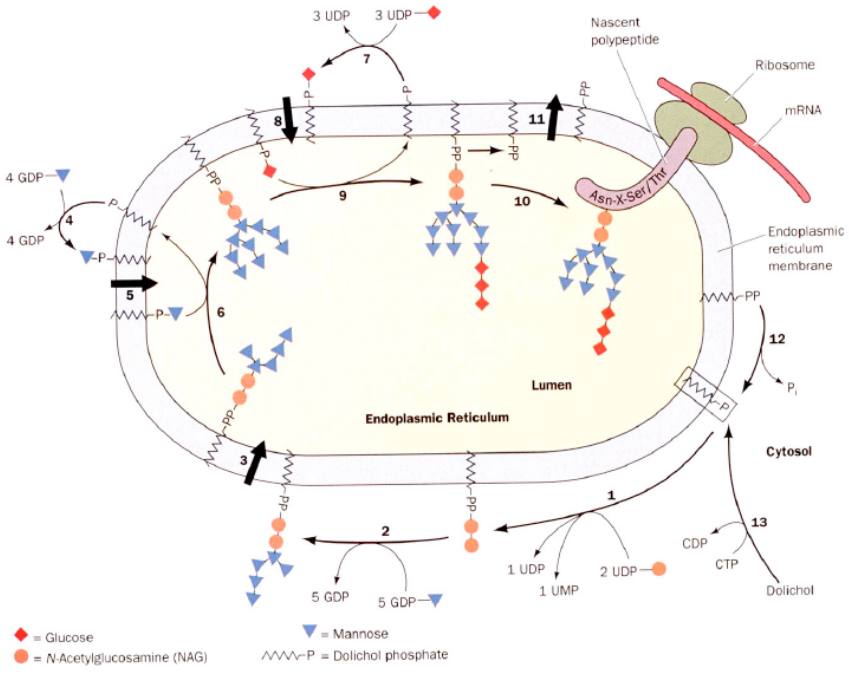
Lecture #33– Protein Glycosylation II

NOTE: This NTC is meant to be used as a study aid to supplement your own class notes. Hence, not all of the text or diagrams contained in the lecture slides will be reproduced here.

Please send any comments or questions about NTCs to either Misha or James through e-mail at **bugs_u2reps@yahoo.ca**

Last lecture we discussed N-glycan synthesis, beginning with the synthesis of the precursor dolichol as a pyrophosphate derivative which occurs in the ER. Using a CTP-dependant kinase reaction, the process begins with the formation of dolichol phosphate. There is a transfer first of 1 N-acetyl glucosamine phosphate with the release of UMP to form dolichol pyrophosphate N –acetyl-glucosamine and then another residue of GlcNAc is added again from the same nucleotide sugar, but this time it is the transfer of a single sugar.

1. GlcNAcs link pyrophosphate to a dolichol. Then there are 5 mannose residues that are added from GDP-mannose, giving a particular mannose-5 structure. The organization of the 5 mannoses is different form that of another structure of 5 mannoses that we will see later on in the same pathway. All of these reactions occur on the cytosolic side of the ER. The nucleotide subunits are produced in the cytosol so that they are readily available for these glycosylation reactions.
2. Following this, there is a flipping of the oligosaccharide lipid precursors so that the carbohydrate now faces the lumen of the ER.
    a. It has been postulated that there are proteins called flippases involved in the translocation. Although their nature has yet to be detailed. WE do know that the flipping is not spontaneous and requires a protein intervention.
    b. All subsequent reactions occur in the lumen of the ER.
3. Since the nucleotide sugars are produced in the cytosol, a topology problem arises: How do the reactions in the lumen utilize sugars and nucleotide sugars which are typically made in the cytosol
    c. The additional mannose residue that are added, are added from another lipid intermediate: dolichol phosphate mannose. Note here it is just a SINGLE phosphate! This dolichol phosphate mannose is made on the cytosolic side of the ER form GDP mannose and dolichol phosphate. Again there must be a flipping of the dolichol phosphate mannose, in order for it to act as a luminal donor of mannose in the ER.
        i. This flippase has yet to be detailed
4. Subsequently, similar reactions add glucose to the oligosaccharide; 3 glucoses are added. There is a formation of a dolichol phosphate glucose from UDP glucose, which again must FLIP ACROSS to the luminal side on order to become a glucose donor



2 Types of reactions occur in the ER

1. Glycosyl transferase reaction: directly transfer sugars from nucleotides
2. Lumenal glycosyl transferase reaction: transfer sugars from a dolichol phosphate mannose/glucose.

The use of these 2 separate reactions solves the topology problem faced during synthesis in the ER lumen.

- The sequence of these reactions and the detail is not seen in the diagram, although the sequence of addition of single sugars is very ordered and specific. Each step is catalyzed by specific enzymes, which makes the formation of the precursor very complex.
- Following these reactions, the dolichol phosphate must flip back to the cytosolic side in order to start the cycle over again. Scientists are trying to elucidate the nature of these reactions.
- Once this precursor has been formed it is transferred from the dolichol pyrophosphate directly to nascent polypeptide chains as they are extruded into the ER (from the translocon) during their synthesis in the cytosol.
- There is an enzyme; oligo saccharyl **transferase**, which carries out this reaction: it transfers the oligosaccharide precursor to Asp X Serine Threonine motifs as they are revealed during the extrusion of the protein chain into the lumen of the ER.
- The **glycosylation** of proteins in Eukaryotes is technically **post translational** but occurs during the extrusion/synthesis of the polypeptide into the ER. The glycosylation of a protein is sequence specific.
- During extrusion into the ER, sites are made available for glycosylation, whereas others are not available due to the folding of a protein. Thus **not all sequences of Asparagine-X-Serine-Threonine are necessarily glycosylated**. Glycosylation of these residue is dependant on their location in the peptide with respect to other given peptide tertiary structures and folding.

The oligo saccharyl transferase is composed of 9 subunits. Its job is to perform 3 tasks:

1. catalytic activity to transfer the oligosaccharide
2. recognize the Asp-X-Ser-Thre sequence
3. recognize specific oligo saccharides

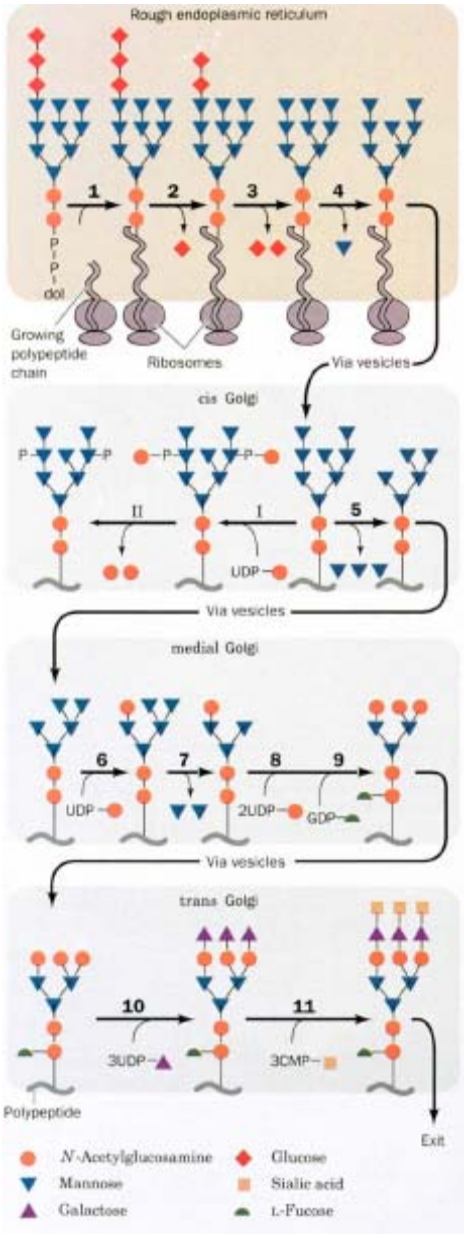
- The elucidation of each of the subunits activities was done by comparing homology of the protein in various organisms throughout evolution. Some more primitive organisms contain only a particular subunit of the enzyme, which they use to produce N-glycans.
- The **oligo saccharyl transferase prefers** oligosaccharides which contain **3 glucose residues**. If there are blocks in the assembly of the oligosaccharide, the enzyme is still capable of transferring these incomplete chains, although its efficiency of transfer is greatly reduced. **Hypoglycosylation** is the resulting phenotype; certain genetic diseases are thought to be related to this event.

Congenital Diseases of Hypoglycosylation

- Some of the enzymes in this pathway and in later parts of the N-glycan pathway contain genetic defects rendering them incapable of efficient glycosylation.
- Individuals with these congenital defects show serious symptoms
  - They are usually mentally retarded and show slow development.
  - N-glycosylation is thus thought to be crucial for normal human development
- Other diseases linked to hypo N-glycan synthesis are being discovered.
- A **defect in the synthesis of GDP mannose** leads to **CDG** (congenital disease of glycosylation) which does not show mental development problems but GI tract problems.
- In vitro studies show that since this is a defect in the synthesis of GDP mannose, these **patients benefit from the consumption of mannose**. This tissue culture study yielded a useful solution to the problem.
- Once this oligosaccharide has been transferred to the polypeptide, the polypeptide must acquire its final conformation which is also dependant on its glycosylation.

Processing of N-glycans; Maturation of N-glycans

- Oligosaccharides with 3 glucose, 9 mannose and 2 GlcNAcs are not the final product found on mature glycoproteins.
- An earlier diagram from the textbook shows the glycoprotein attachment to the polypeptide (extruded by ribosome) much too long after its synthesis. This attachment occurs early on in protein extrusion.
- The transfer of the oligosaccharide to nascent peptides occurs first with the removal of the terminal glucose residue, which can occur translationally, while the polypeptide is still bound to the ribosome.
- The **oligosaccharyl transferase and the glucosidase I are located very near the translocon.**
- Subsequently when the polypeptide is released, glucosidase II removes the other 2 glucose residues. This yields a glucose free oligosaccharide. A mannosidase in the ER produces an oligo saccharide with 8 mannose residues with a specific arrangement of the mannose residues on the chain.
- Isomers of this structure are constant; the formation of these residues is not random.
- All of the reactions which occur in the ER are **highly conserved in most eukaryotes**; from yeast to mammals. Some trypanosomes have different mechanisms, but for our purposes we will not be discussing them.



Onward to the Golgi!

- Once this **mannose 8 structure** is formed on the newly formed glycoproteins, the glycoproteins are transported to the Golgi. As they are **transported through the Golgi**, various reactions can occur.
- If we follow reaction #5 which occurs in the Golgi, we see that 3 different **mannosidases called 1a, 1b and 1c remove 3 additional mannose** residues to yield a structure with **5 mannoses**.

The arrangement of the mannoses during the synthesis (in the ER) of the dolichol oligosaccharide is very different to that of the man-5 structure during the processing in the Golgi of the oligosaccharide

- This reveals why there is microheterogeneity.
  - If the glycoprotein is moving quickly through the Golgi, the relative activity of all the enzymes will be affected producing different combos of mannose structure**.**
- **His 5-man structure is the substrate for the 1st glycosyl transferase** (GlcNAc transferase 1) in the Golgi, which **adds a single residue of GlcNAc** to a particular branch of the oligosaccharide which initiates a branch of complex saccharides.
  - This enzyme is essential for the initiation of the formation of complex hybrid N-glycans.
- One the 1st GlcNAc has been added, 2 other enzymes; mannose Golgi mannosidase 2 and 2X (similar) are responsible for the removal of another 2 mannose residues.
- If the Golgi mannosidase 2 is inhibited or non functional, hybrid structures are NOT produced, since the final 2 mannoses are not removed.
- To form complexes, these 2 mannoses must be removed to uncover another mannose which becomes yet another substrate from other trasnferases. **GNT1, 2, 3, 4, 5 & 6 can produce highly branched structures, using other sugars.**

- Bi-ternary, tri-ternary and tetra-ternary structures are possible. **Which all require the function of GNT1.**
  - Golgi function is dependant on the cell type and species. These reactions are not as conserved as those in the ER.
- This assembly line process occurs during movement of the glycoproteins through the Golgi, which is flexible and produces microheterogeneity.
  - The relative concentrations of enzymes, the rate of glycoprotein migration and the efficiency of each enzyme all have an impact on the final glycoprotein structure.
- Since nucleotide sugars are made in the cytosol, and they are utilized in the lumen of the Golgi, a topology problem presents itself.
  - This situation is different from the ER; there are **NO lipid intermediate and no dolichol derivatives**! In fact these sugars are transported to the Golgi lumen using specific **nucleotide sugar transporters.**

Mice Knockouts & the elucidation of carbohydrate function

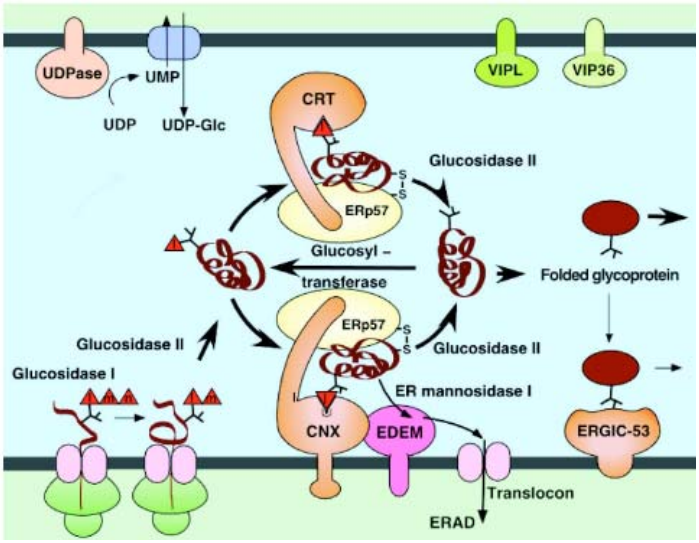
- Many mice knockouts have been created to observe the role of carbohydrates during development and various types of diseases.

For example:

- The 1st knockout created got rid of **GNT1** (reaction #6) which initiated formation of complex hybrid N-glycans. This knockout is **embryonic lethal**, the mouse will never be born. This shows that the formation of these complex hybrid oligosaccharides is crucial for development.
- At the beginning of this year, a **knockout of 1 out of 3 mannosidase** genes was created (these transform the oligo saccharide to man-5). The formation of this man-5 is essential for NGT1 function (since it is the substrate).
  - Since there are 3 genes, the phenotypes of each should be different from that of a gene knockout for the GNT1.
  - Mice are **born normally**, but after a few hours they go into **respiratory distress** and die of massive lung haemorrhage.
  - 3 enzymes carry out the same reaction, but they have different distributions in various cell types. The knockout gene was assumed to be essential for the later stages of development, since the mice are born but do not have fully developed lungs.
- As one goes down the pathway and knocks out later enzymes such as galactosyl transferase, the result is not an embryonic lethal phenotype.
  - The defects become more specific and become more subtle. Their roles become apparent as a result of pathophyisiology but these knockouts are not crucial to the development of organ systems.
- Pathways numbered I and II (diagram on previous page) are diversions form the standard pathway of glycosylation.
  - These pathways are specific for lysosomal enzymes. Many digestive enzymes exist in lysosomes, which must be targeted to the lysosomes through glycosylation (major mechanism in animal cells).
- An addition of GlcNAc phosphate (from UDP GlcNAc) in several positions of the high mannose oligosaccharide is required for targeting to lysosomes.
  - N-acetylglucosaminyl phospho**transferase** (in Golgi) adds GlcNAc-P to Mannose.
    - This **enzyme is specific for lysosomal proteins** and does not act on other glycoproteins. This enzyme recognizes a conformational epitope only found on lysosomal "destined" proteins, it is not a specific sequence but a 3-D structure of the protein. Genetic defects in this enzyme and thus loss of targeting results in secretion of lysosomal enzymes (I-cell disease).

- N-acetylglucosamine-1-phosphodiester α-Nacetylglucosaminidase (**uncovering enzyme**) removes GlcNAc and exposes Man-6-P.
- Man-6-P is a recognition marker for specific lectins present in the Golgi and form a complex allowing the lysosomal enzymes to be transported into the lysosome.
- The reactions in the Golgi (N-glycans as well as other form of glycosylation) are different from those of the ER.

Role of carbohydrates in glycoprotein folding



- The folding of a polypeptide is dependant on its sequence. Certain proteins require the assistance of chaperones, while others fold unto themselves spontaneously. Folding also involves the formation of disulfide bonds between amino acids particularly for secretory proteins.
- Glycoprotein folding is very efficient and is termed the **"calnexin cycle" which involves 2 proteins**
  - membrane bound **calnexin**; preference for membrane glycoproteins
  - soluble **calreticulin**: preference for soluble glycoproteins
- These 2 proteins perform similar function although they have preferences for certain glycoproteins.
- The initial stages of processing begin with the removal of 1 glucose residue by glucosidaseI.
- Then there is removal of an additional glucose residue by glucosidase II.
  - At this stage some glycoproteins may not have yet acquired their final conformations. For more complex glycoproteins which do not spontaneously fold, an incompletely folded peptide with 1 remaining residue on the carbohydrate (**mono-glycosylated**) will be recognized by calnexin or calreticulin.
- **Calnexin and calreticulin are both lectins** because they recognize carbohydrates. They are found as part of a complex with the protein **Erp57** which catalyzes the formation of proper disulfide bonds.
- Glycoproteins will be released from this complex, and are acted upon by glucosidase II which will remove a 2nd glucose residue.
  - If at this stage, the glycoprotein has yet to acquire its final conformation, there is a specific soluble **glucosyltransferase**, found in the ER lumen which adds glucose from UDP-glucose back onto the unfolded glycoprotein. This enzyme is a **sensor of incompletely folded glycoproteins**; it does not work on native glycoproteins. It allows glycoproteins to go through the folding cycle again.
  - 2 domains:
    - **catalytic domain**: adds glucose (from UDP glucose transported into the ER)
    - **Domain capable of recognizing hydrophobic patches** on improperly or incompletely folded glycoproteins.
- Some glycoproteins may never acquire their final conformation, either randomly or due to errors in translation. These proteins must be targeted for degradation, which is part of quality control.
- The cell will determine which proteins are retained and those that must be degraded. Quality control of glycoproteins is dependant on **mannose trimming**.
  - Somehow this trimming done by the ER mannosidase (and others) targets the glycoprotein for degradation. They are exported out of the ER through a translocon and are transported into the cytosol through a process called **ER associated degradation (ERAD)**

- **ERGIC-53** will recognize high mannose oligosaccharides on certain glycoproteins in order to facilitate their export out of the ER.
  - Homologs of ERGIC-53 have been discovered although not much is know about the specific nature of their function, they are most likely involved in intracellular targeting.

O-linked oligosaccharide structure and biosynthesis of mucin-type (GalNAc-Ser/Thr)

- N-glycan is the most complicated synthetic scheme, although others exist.
- GlcNAc attached to serine and threonine is a form of O-glycan, which occurs on both soluble and membrane glycoproteins and does not contain a specific sequence for the addition of O-glycan.
  - They are synthesized in the golgi through the stepwise addition of each sugar (from nucleotide sugar). This process is relatively simple in comparison to N-glycan synthesis.
- The core structures found on these O-glycans contains GlcNAc linked to serine or threonine, but they may **acquire linear or branched conformations**.
- Mucins are a form of this type of protein, and are highly glycosylated (as high as 90% carbohydrate).
  - **Mucin is made up of tandem repeats of amino acids** (rich in serine and threonine), with **extended conformation** (not globular) and resemble rods. Mucins are the major component of mucus lining all epithelia cells, particularly in respiratory tract.
  - It is extremely important since it imparts viscosity to the mucus which acts as a protective barrier of the epithelium. When mucin is damaged there is increased risk of infection.
- Mucins contain a core structure as well as a backbone structure (either linear or branched) and peripheral/terminal sugars; which may contain a variety of specificities.

O-glycan specificities in Blood groups

- Blood groups cannot be transfused, this will cause anaphylactic shock.
- The A, B and O blood groups are determined by specific structures of mucin type glycoproteins as well as other glycoproteins. The linkage of the repeating backbone will yield either type I or type II (Gal β 1-3 GlcNAc or Gal β 1-4 GlcNAc).
- The terminal sugars are either
  - fucose α1-2 (present in blood group O with H specificity)
  - GlcNAc α 1-3 linked to precursor structure yield A blood group
  - Gal α 1-3 linked to precursor structure yield B blood group.
- Blood group O contains no transferase enzymes.
- Blood group A utilizes a specific GlcNAc transferase which adds it to the precursor whereas in blood group B a specific Gal transferase acts on the peptide.
- 4 amino acids separate these 2 enzymes, meaning that they are very similar.
- Blood group AB, each allele encodes a different transferase.

Summary of Carbohydrate Functions →

Structures responsible for A, B, H specificities

Specificity	Structure: Type 1	Structure: Type 2
H (O)	Gal-β-(1→3)-GlcNAc.. 2 ↑ L-Fuc-α-1	Gal-β-(1→4)-GlcNAc.. 2 ↑ L-Fuc-α-1
A	GalNAc-α-(1→3)-Gal-β-(1→3)-GlcNAc.. 2 ↑ L-Fuc-α-1	GalNAc-α-(1→3)-Gal-β-(1→4)-GlcNAc.. 2 ↑ L-Fuc-α-1
B	Gal-α-(1→3)-Gal-β-(1→3)-GlcNAc.. 2 ↑ L-Fuc-α-1	Gal-α-(1→3)-Gal-β-(1→4)-GlcNAc.. 2 ↑ L-Fuc-α-1

STRUCTURAL ROLE
- ◆ CONFORMATION
  - protein folding – calnexin cycle
  - quality control in the ER -ERAD
- ◆ STABILITY
  - resistance to proteolysis
- ◆ PHYSICAL PROPERTIES
  - increased solubility
  - increased viscosity – e.g. mucins

EXAMPLES OF RECOGNITION FUNCTIONS
- ◆ ANTIGENIC DETERMINANTS
  - ABO blood groups
  - (onco)developmental antigens
- ◆ ADHESION OF PATHOGENS
  - bacteria, viruses, toxins
- ◆ CELL ADHESION
  - recruitment of leukocytes by selectins
- ◆ CLEARANCE FROM CIRCULATION
  - pituitary peptide hormones, e.g. LH
- ◆ INTRACELLULAR TARGETING
  - lysosomal enzymes

SIGNAL TRANSDUCTION
- ◆ MODULATION OF RECEPTOR SIGNALLING
  - Notch receptor - O-linked fucose on EGF repeats