

Chapter 3

The Major Glycosylation Pathways of Mammalian Membranes

A Summary

Ajit Varki and Hudson H. Freeze

1. INTRODUCTION

Many types of cellular membranes have sugar chains (*oligosaccharides* or *glycans*) attached to a variety of their constituent macromolecules. Such glycosylated macromolecules are called *glycoconjugates*. This chapter will broadly scan the subject of structure and biosynthesis of oligosaccharides on glycoconjugates, primarily in higher animal cells. The classification of different types of oligosaccharides will be based on the type of *linkage* that joins them to the nonsugar component (*aglycone*), rather than on the types of sugars (*monosaccharides*) found in the chains. As shown in Table I, this classification is slightly different from the more conventional categories and nomenclature used for glycoconjugates in the past. This change is made because it is now clear that the aglycone usually determines the initiation step in creating a glycoconjugate and can have an influence upon the further buildup of a more complex glycan. Furthermore, the traditional boundaries between the different types of glycoconjugates as previously described have become blurred.

Ajit Varki and Hudson H. Freeze Cancer Center, University of California at San Diego, and
La Jolla Cancer Research Foundation, La Jolla, California 92093.

Subcellular Biochemistry, Volume 22: Membrane Biogenesis, edited by A. H. Maddy and J. R.
Harris. Plenum Press, New York, 1994.

Table I
Descriptive Terms for Glyconjugates in Relation to the Attachment Site
of Sugar Chains

Historical terms	Conventional terms	Defining types of linkages and sequences
Glycoprotein	<i>N</i> - or <i>O</i> -linked glycoprotein	Scattered <i>N</i> -GlcNAc and/or <i>O</i> -GalNAc-linked sugar chains
Mucins Mucopolysaccharides	Mucin-like glycoproteins Proteoglycans	Many-clustered <i>O</i> -GalNAc-linked chains <i>O</i> -Xylose-linked glycosaminoglycan chains on core proteins Can also carry <i>N</i> -GlcNAc and/or <i>O</i> -GalNAc-linked chains
Mucolipid, glycolipid Not recognized earlier	Glycosphingolipid Glycophospholipid anchored protein	Glucose or galactose linked to ceramide Ethanolamine-phosphate in amide-linkage to proteins and glycan linked to phosphatidylinositol

The uniformity of the glycone-aglycone linkage region and of the first few sugars immediately attached to it forms a sharp dividing line between the different classes of oligosaccharides. However, this distinction becomes less prominent as more monosaccharide units are added. In fact, considerable overlap is seen in the structures of outer sequences of the different types of glycans. It is also important to remember that an individual aglycone (e.g., a protein) can carry many different types of oligosaccharide chains attached by different linkages. Most types of oligosaccharides are synthesized in the membrane-bound compartments of the ER-Golgi-plasmalemma pathway (see Figure 1). Thus, trimming and restructuring reactions in the biosynthesis of a variety of sugar chains can occur in parallel on a single glycoconjugate during its synthesis, trafficking, and maturation.

1.1. Historical Overview

The study of carbohydrate-containing molecules began with the analysis of specific classes of macromolecules. Each such class was in easily recognizable forms that could be defined and separated by their physical properties (e.g., viscous secretions, soluble blood proteins, amphipathic lipids, and gristly connective tissue molecules). The carbohydrate analytical technology of the time required large amounts of material and allowed somewhat gross analyses of these apparently different types of molecules. Thus, individual research groups tended to focus upon analysis of different types of glycosylated molecules such as glycoproteins (with *N*-linked oligosaccharides), mucins (with *O*-linked oligosaccharides), glycolipids (with lipid tails), and proteoglycans (with glycosaminoglycan chains) (see Table I). Both the diversity and complexity within each of these

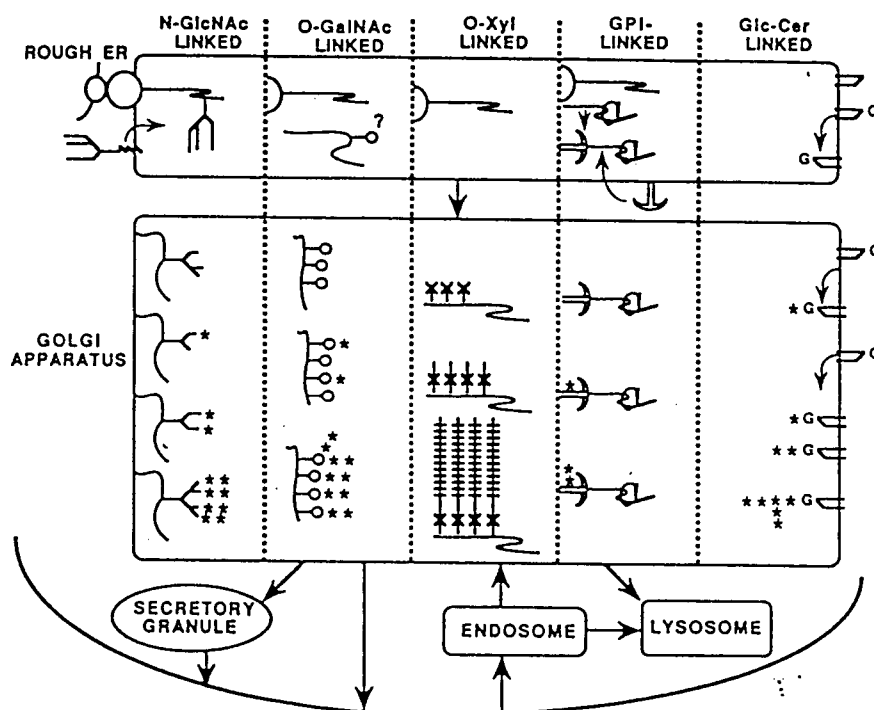


FIGURE 1. Initiation and maturation of the major types of glycoconjugates in relation to subcellular trafficking in the ER-Golgi-plasmalemma pathway. The asterisks represent the addition of outer sugars to the oligosaccharides in the Golgi apparatus. See text for a better understanding of the events depicted in the figure.

individual groups of molecules were challenging, and the major structures appeared to be clearly different in each group. Thus, research groups seldom crossed over from studying one type of molecule to another. For the same reason, prior reviews of the structure, biosynthesis, and function of glycoconjugates have tended to focus upon one or the other of such specific classes of molecules.

1.2. The Modern View

More refined and sensitive technology and the elucidation of biosynthetic pathways made it possible to analyze trace amounts of functionally important molecules (Cummings *et al.*, 1989; Varki, 1991). Such studies showed that the dividing lines between the different glycoconjugates are not as rigid as previously thought. Thus, some glycoproteins contain functionally important glycolipid portions, some glycoproteins carry glycosaminoglycan chains, and many proteoglycans have both *N*- and *O*-linked carbohydrate chains. Also, the same di- or tri-

saccharide unit is frequently found in the outer chains of both glycoproteins and glycolipids. Such combinations occur on many important molecules, and in some cases the glycans themselves determine function. Thus, the traditional lines of distinction have become blurred. Focus on the composition of each general class of glycoconjugate is now being replaced by analysis of individual sugar chains on physiologically important molecules. The novel term "glycobiology" describes the study of the structure, biosynthesis, distribution, regulation, and function of the entire spectrum of sugar chains found on glycoconjugates (Rademacher *et al.*, 1988).

In this chapter we will survey the major types of sugar chains and their biosynthesis. This includes the minimal requirements to initiate their synthesis, how the chains are extended, terminated, and turned over. We will also mention the subcellular localization of these biosynthetic steps in the general ER-Golgi-plasmalemma pathway, and some inhibitors and mutants useful in their study (Elbein, 1991; Winchester and Fleet, 1992). As discussed above, we will classify these sugar chains by their unique linkage regions. Because of the broad nature of this subject, the references cited consist only of review articles. For the same reason, this chapter does not deal with the function of these oligosaccharide chains, a subject deserving of a completely separate review (Cumming, 1991; Stanley, 1992; Varki, 1992b; Varki, 1993).

2. THE ER-GOLGI-PLASMALEMMA PATHWAY

The vast majority of glycoconjugates follow a common trafficking pathway involving biosynthesis in the lumen of the ER and passage through the Golgi apparatus (Farquhar, 1991; Mellman and Simons, 1992; Rothman and Orci, 1992; Chapter 4 this volume). They may then remain membrane bound, be secreted immediately, or be packaged into secretory granules or lysosomes (see Figure 1). While some bulky sugar chains are made on the cytoplasmic face of these intracellular membranes and "flipped" across to the other side, most are added to the growing chain on the *inside* of the ER or the Golgi apparatus (Hirschberg and Snider, 1987). An important topological point to remember is that whatever portion of a molecule faces the *inside* of the ER will ultimately face the *inside* of a secretory granule or lysosome, but it will face the *outside* of the cell. This is the topology of most forms of glycosylation. The biosynthetic enzymes (glycosyltransferases) responsible for these reactions are well studied, and their location can help define various functional compartments of the ER-Golgi pathway (Paulson and Colley, 1989). A working model envisions the enzymes as physically lined up along the pathway in the sequence in which they work. This is probably an oversimplified view, and the distribution probably depends upon the cell type. The low-molecular-weight sugar nucleotides that act as donors for most of the biosynthetic steps are made in the cytosol, and specifically transported into the lumen of the organelles. A major exception to this rule

is the addition of single GlcNAc residues to cytoplasmic proteins; the active site of this transferase faces the cytosol. There are also other types of cytoplasmic glycosylation, but the important point is that they tend to be very different from those found in the ER-Golgi-plasmalemma pathway.

In some cases, the oligosaccharide components of glycoconjugates (particularly the outer sugar units) appear to be turned over more rapidly than the aglycone that carries them. This implies considerable metabolic activity of these molecules and their recycling back into the Golgi for reglycosylation. The purpose of such turnover and resynthesis is currently unknown.

In the following sections, the different types of linkage regions found in the ER-Golgi-plasmalemma pathway are individually considered, with regard to their initiation signal, donor molecules, processing and termination, and specific inhibitors and mutants in their biosynthesis. The text presents limited information on the major principles, but many additional facts can be found in the accompanying figures and tables. Consult the references for more comprehensive, detailed information.

2.1. The Generation of Diversity

Animal oligosaccharides contain a limited set of monosaccharides and generate diversity by using different linkages. Of the dozens of possible monosaccharides, only a few are actually used in mammalian cell glycoconjugates (Kisailus and Allen, 1992). Figure 2 shows these sugars, along with abbrevia-

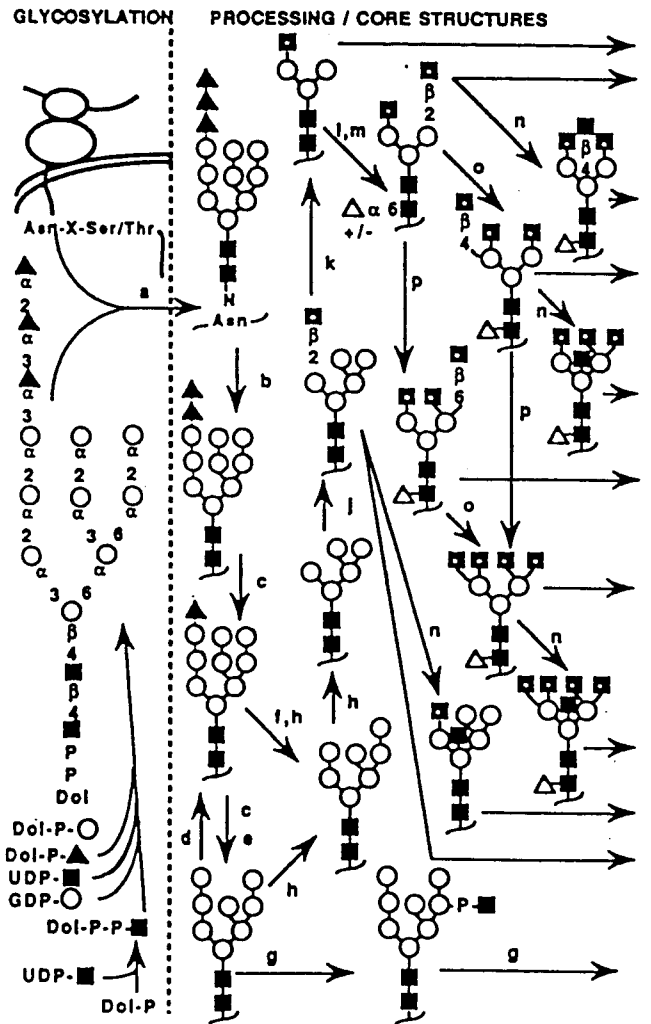
▲ = Glucose (Glc)	△ = Fucose (Fuc)	Ac = O-acetyl
○ = Mannose (Man)	▽ = Xylose (Xyl)	P = Phosphate
● = Galactose (Gal)	◇ = Sialic acids (Sia)*	S = O-Sulfate
■ = N-acetylglucosamine (GlcNAc)	◇ = Glucuronic acid (GlcA)	NS = N-Sulfate
□ = N-acetylgalactosamine (GalNAc)	◆ = Iduronic acid (IdA)	NH ₂ = free amino group

- All monosaccharides are in the D-configuration except for Fucose and Iduronic acid, which are in the L-configuration
- All monosaccharides linkages originate from the 1-position except for the sialic acids, which are linked from the 2-position
- All glycosidically linked monosaccharides are in the pyranose form

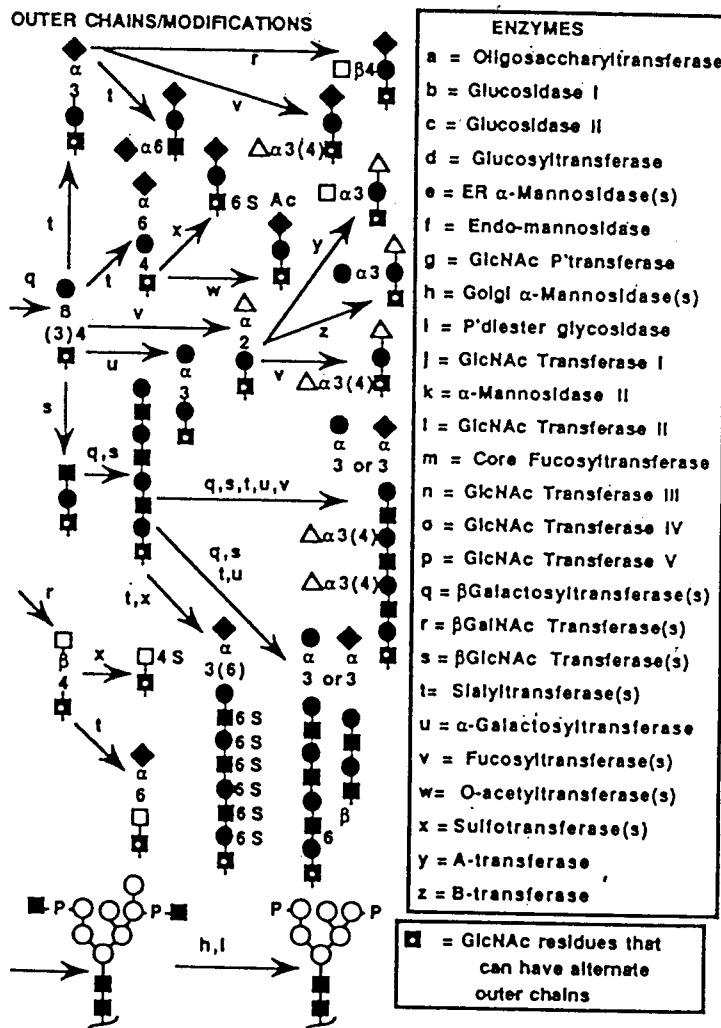
*Common name for a family of nine-carbon acidic monosaccharides

SAMPLE: 9OAc-Siaα2-6Galβ1-4GlcNAcβ1- = 9Ac ◆ α6 ● β4 ■ β1-

FIGURE 2. The monosaccharide units and substituents of mammalian oligosaccharide chains. Monosaccharides used in the biosynthetic pathways described in this chapter are presented, along with their standard abbreviations in parentheses. The symbol key for the monosaccharides is used throughout the remainder of the figures. Because the origin of a linkage is invariant for each monosaccharide, the remaining figures only indicate the position of attachment to an underlying sugar.

FIGURE 3. The *N*-GlcNAc-linked oligosaccharide

tions, and a symbol key that will be used for all of the figures in this chapter. Each monosaccharide can be attached in α - or β -linkage from its 1-position (2-position only for sialic acids) to an underlying acceptor. If that underlying acceptor is another monosaccharide, the next number indicates the position on the acceptor where the attachment occurs (e.g., Gal β 1-4GlcNAc indicates a galactose residue β -linked via its 1-position to the 4-position of an *N*-acetylglucosamine residue). Because the first number is invariant for each monosaccharide, it is not indicated in the figures. Various modifications of the monosaccharides can also occur, and



pathway. See Figure 2 for symbol key.

these are listed in Figure 2. The sialic acids designated by "Sia" actually represent a family of nine-carbon acidic sugars (Varki, 1992a; Schärer, 1991).

Literally hundreds of glycosyltransferase enzymes must be involved in the biosynthesis of the different types of oligosaccharides discussed in this chapter. No details regarding these enzymes will be presented here. Several of them have been cloned recently (Paulson and Colley, 1989). Surprisingly, while these enzymes carry out similar reactions, using similar donors and acceptors, sequence homology

with one another is uncommon. However, they all show a common structural motif, being type-2 membrane-bound proteins with short cytosolic amino-terminal sequences, followed by hydrophobic membrane-anchor domains and globular carboxylterminal domains carrying the catalytic site within the Golgi lumen.

2.2. *N*-GlcNAc-Linked Oligosaccharides

These oligosaccharides are defined by the GlcNAc-*N*-Asn linkage to proteins (see Figure 3). A preformed precursor glycan is transferred from a lipid donor to the acceptor protein on the luminal side of the ER, either co-translationally or immediately posttranslationally, to form a core structure that is subsequently processed (Kornfeld and Kornfeld, 1985; Cummings, 1992; Kobata and Takasaki, 1992; Schachter, 1991).

2.2.1. Initiation

The consensus sequence dictating the addition of these oligosaccharides is Asn-X-Ser/Thr (where X can be any amino acid except Pro). This simple tripeptide (with blocked amino and carboxy termini) can alone act as an acceptor when fed to intact cells or when mixed with isolated ER membranes. However, not all such sites are utilized in natural proteins. The structure of the protein and the rate of its synthesis can apparently influence whether a potential glycosylation site is actually used.

2.2.2. Donor

The full-sized lipid-linked oligosaccharide (LLO) is synthesized in the ER on an isoprenyl lipid called *dolichol phosphate*, by a series of sequential glycosyl transferase reactions. The first step in this synthesis is blocked by the antibiotic tunicamycin (Table II). Truncated forms of the LLO synthesized in mutant cells or those made in the presence of inhibitors can also be transferred to proteins, but usually not as efficiently as the full-sized sugar chain. Although the LLO is initiated on the outside of the ER, its completion occurs on the inner (luminal) surface of the ER where it is transferred to the acceptor. The mechanism used to flip the bulky, hydrophilic oligosaccharide through the membrane is not known. Any mutations in the synthesis of the LLO will affect *all* proteins with *N*-GlcNAc-linked chains.

2.2.3. Processing and Elongation

Figure 3 outlines the complex and variable processing of *N*-linked glycans. The actual pathway followed for an individual glycosylation site on a given protein is determined by many factors including the structure of the protein, the location (availability) of the sugar chain on the protein, the relative activities of

Table I
The Integrin Family and Adhesion Receptors

Adhesion molecules		Receptors	
Names	Regulation of surface expression	Names	Regulation of surface expression
1. The β_1 integrins involved in cell-extracellular matrix adhesion			
$\alpha_1\beta_1$ (VLA-1)	Constitutively activated	Collagen	Extracellular matrix proteins
$\alpha_2\beta_1$ (VLA-2)		Collagen types I-IV, VI	
$\alpha_3\beta_1$ (VLA-3)		Fibronectin, laminin, collagen	
$\alpha_4\beta_1$ (VLA-4)		Fibronectin	
$\alpha_5\beta_1$ (VLA-5, fibronectin receptor, GP ^{IIb/IIIa})		Fibronectin	
$\alpha_6\beta_1$ (VLA-6)		Laminin	
2. The β_2 integrins involved in leukocyte-endothelial cell adhesion			
LFA-1 or CD11a/CD18 or $\alpha_L\beta_2$ integrin	C5a, PAF Constitutively present on the plasma membrane and translocated to it by upregulation	Counter-receptors ICAM-1 or CD54 ICAM-2	IL1, TNF α , slow response
Mac1 or Mol or CD11b/CD18 or OKM1 or $\alpha_M\beta_2$ integrin	C5a, fMLP, IL8, PAF, LTB $_4$, fast response	ICAM, bound iC3b	IL1, TNF α , fast response

(continued)

Table III
Examples of Glycosylation-Defective Tissue Culture Cell Lines

	Parental line	Mutant line	Altered step	Mature carbohydrates	
Single pathway	BW5147	PHA ^R 2.7	α -Glucosidase II	Glycoproteins with Glc ₂ Man ₉ GlcNAc ₂ . Complex chains decreased.	
	L cells	CL6	α -Mannosidase I	Glycoproteins with Man ₉ GlcNAc ₂ . Complex chains decreased.	
	CHO	Clone 15B Lec 1	GlcNAc transferase I	Processing blocked at Man ₃ GlcNAc ₂ processing.	
	BW5147	PHA ^R 1.8	GlcNAc transferase V	Decreased multiantennary N-linked oligosaccharides.	
	CHO		745	Xylosyl-transferase	No glycosaminoglycan chains.
			761	Gal-transferase-I	No glycosaminoglycan chains.
	Human fibroblasts	I-cell disease	Phosphorylation of lysosomal enzymes	Lack of M-6-P residues.	
Multiple pathways	CHO	Clone 13, Lec 8	UDP-Gal-transport into Golgi	Glycoproteins/glycolipids decreased in Gal and Sia residues. Complex glycans terminate in GlcNAc.	
	CHO	Clone 1021 Lec 2	CMP-Sia transport into Golgi	Glycoproteins/glycolipids terminate in Gal and have decreased sialic acid.	
	CHO	<i>Id1^a</i>	UDP-Gal/GalNAc 4'-epimerase	Deficient in Gal and GalNAc-residues. Affects O-GalNAc initiation, and addition of Gal to N-GlcNAc linked and ceramide-linked molecules.	

the various transferases and processing glycosidases, and the transit rate through the Golgi apparatus. The early processing steps occur in essentially all organisms, plant and animal, but continued processing leads to many diverse structures that are special for higher animals.

All of the Glc residues are removed in the ER, followed by removal of a variable number of Man residues. Metabolic inhibitors and mutant cell lines can halt processing at many of these early steps (see Tables II and III). The addition of the first GlcNAc residue represents the transition from "high-mannose" type to "complex type" oligosaccharides. However "hybrid" structures can also result. The processing of Man residues continues and it is followed by the addition of a variable number of β -GlcNAc, β -Gal, and α -sialic acid residues in the most common pathways. The α -fucose residues can be added in the core region or in the peripheral regions, and in some cases terminal α -Gal residues are substituted for sialic acids. One modification exclusive for *N*-linked oligosaccharides is the addition of Man6P residues to high mannose chains on lysosomal enzymes. These residues function to target the newly synthesized enzymes to the lysosome by binding to a receptor in the Golgi apparatus.

2.2.4. Termination

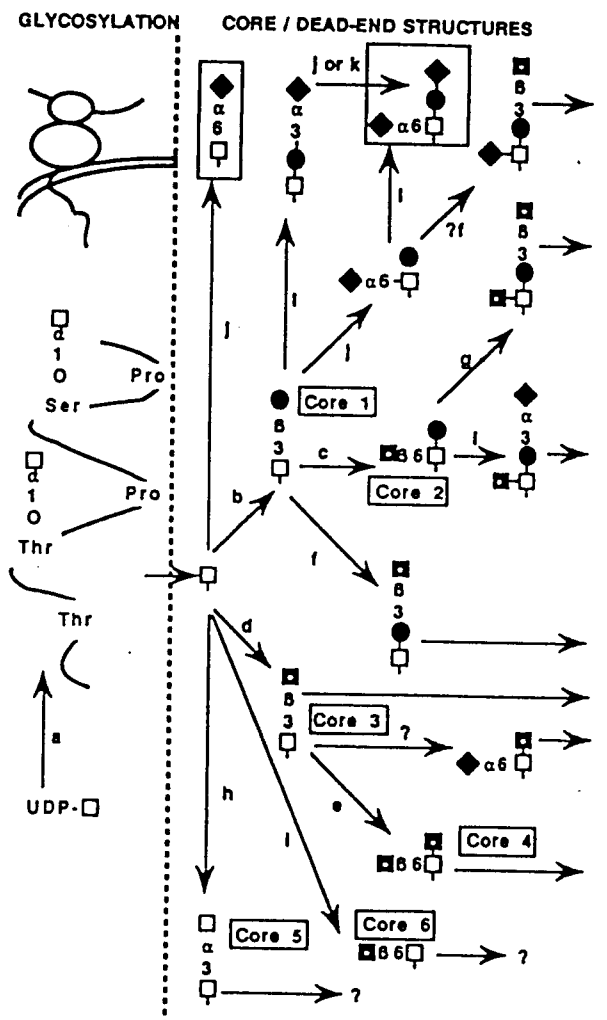
Certain monosaccharide units are typical of the nonreducing ends of many glycoconjugates. Chains can be terminated by α -sialic acids, α -fucose, α -Gal or blood group sugar sequences. After these transferases add their respective monosaccharides, some of the oligosaccharides can be modified further (e.g., by sulfation or acetylation) as a final step in biosynthesis. How all of these terminal reactions compete with one another is an interesting issue that is not well resolved.

2.2.5. Fate and Turnover

Proteins with *N*-linked chains are found in membrane-bound organelles, and as secreted molecules. On some cell-surface proteins, the terminal sugars (e.g., sialic acid) turn over faster than the peptide itself, and can also be re-added after removal. Secreted serum proteins can also lose terminal sugars. In mammals and birds, this loss may serve as a signal for the proteins to be cleared from the circulation by carbohydrate-specific receptors found on liver cells.

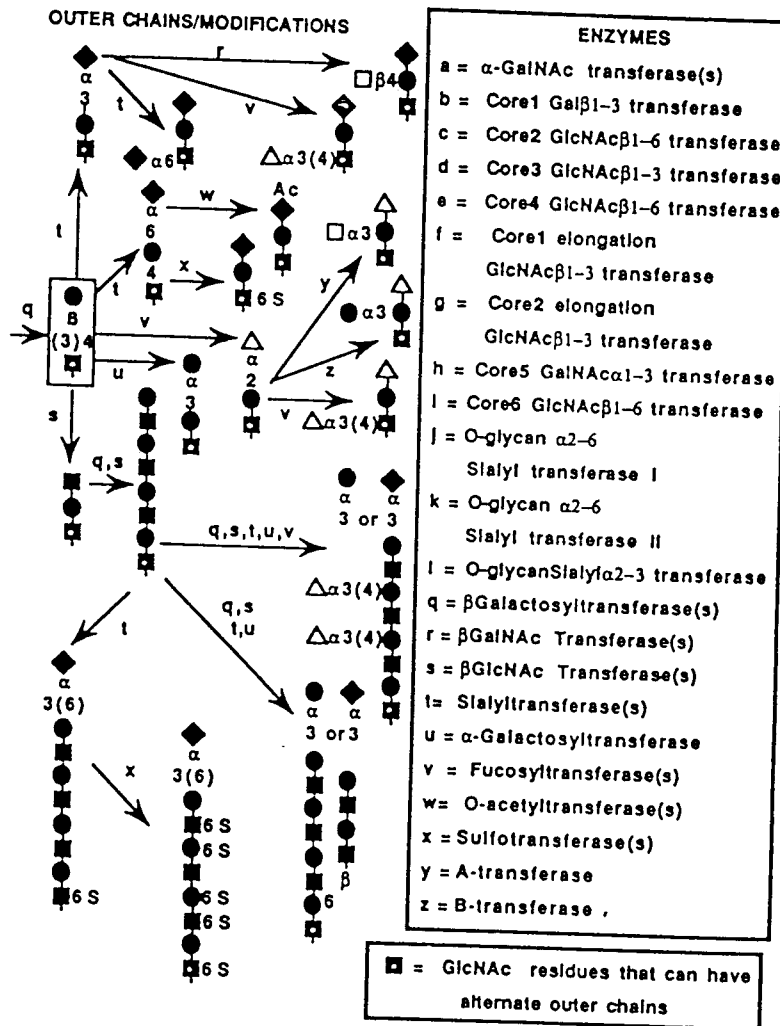
2.3. *O*-GalNAc-Linked Oligosaccharides

These sugar chains are linked to a protein at Ser or Thr residues through an α -GalNAc unit. These chains can vary greatly in size. They are often, but not always, found in clusters on the polypeptide backbone. Unlike *N*-linked oligosaccharides, no lipid-linked precursors or *en bloc* transfers occur (Schachter and Brockhausen, 1992; Carraway and Hull, 1991; Fukuda, 1992).

FIGURE 4. The *O*-GalNAc-linked oligosaccharide

2.3.1. Initiation

There does not appear to be a single consensus sequence for the addition of *O*-linked sugars. However, the peptide will often have a Pro-residue at positions -1 and +3 relative to the glycosylated Ser or Thr residue. Clusters of *O*-GalNAc-linked oligosaccharides are frequently found in regions of the protein with a high density of alanine, serine, threonine, and proline residues. The difference between Ser and Thr acceptors indicates that initiation requires two different transferases, for which there is now evidence. Initial *O*-glycosylation probably occurs predominantly in the Golgi apparatus.



pathway. See Figure 2 for symbol key.

2.3.2. Elongation

Several different pathways are available for the elongation of *O*-linked chains as shown in Figure 4. The early steps in these pathways generate a series of distinctive units called "core" units, numbered 1 to 6. Commitment to one pathway can block access to another pathway leading to different structures. The activity of certain inhibitors (see below) shows that the transferases acting later in this pathway do not need to recognize the protein backbone to elongate the chains.

2.3.3. Termination

The final steps are very much like those seen in the synthesis of *N*-linked chains discussed above.

2.3.4. Inhibitors and Mutants

This type of glycosylation can be partially blocked by adding millimolar concentrations of phenyl or benzyl α -GalNAc to cells. These hydrophobic inhibitors penetrate the cell and the Golgi and compete with the endogenous protein acceptors. The artificial derivatives then transit through the Golgi and can act as acceptors for other sugar residues. In contrast to the *N*-linked oligosaccharide pathway, cellular mutants that selectively affect *O*-linked oligosaccharide biosynthesis are uncommon. However, several pleiotropic mutants affect both *N*- and *O*-linked oligosaccharide biosynthesis (see Table III).

2.3.5. Fate and Turnover

Many of the proteins with clusters of short *O*-linked chains are the traditional "mucins" that are secreted in large quantities on mucosal surfaces. On the other hand, some integral membrane proteins such as the LDL-receptor contain clusters of *O*-linked chains on the extracellular side of the transmembrane region. Such clustered *O*-linked chains tend to extend the polypeptide into a rod-like section. Thus, these clusters may help to prevent proteolysis of the proteins, to prop up the globular domains of receptors away from the membrane, and to present clustered ligands for recognition by receptors.

In at least one case, a protein with a large number of *O*-linked chains has been found to recycle repeatedly between the Golgi and the cell surface. On successive visits to the Golgi, new chains can be initiated and built on previously unmodified amino acids. It is not known whether this is a general property of all membrane-bound proteins with *O*-linked chains. This type of change does not generally happen for *N*-linked chains. This is because the latter are first added in the ER, which is not generally considered to be in the recycling pathway from the surface.

2.4. Ceramide-Linked Oligosaccharides (Glycosphingolipids)

These molecules are defined as the glucosylceramides and the galactosylceramides, and are predominantly, but not exclusively, found on the cell surface. They are also called glycosphingolipids because of the sphingosine chain that makes up part of the ceramide core (Hakomori, 1986, 1990; Kundu, 1992; VanEchten and Sandhoff, 1993; Shayman and Radin, 1991; Stults *et al.*, 1989).

2.4.1. Initiation

The ceramide lipid tail is made in the ER membrane, facing the cytoplasm. The addition of glucose (the first sugar) is believed by most investigators to occur on the cytosolic face of the membranes of the ER or the Golgi. The glucosylceramide is then probably flipped into the lumen for further addition of sugars. Less is known about the initial addition of galactose to the ceramide.

2.4.2. Elongation

Figure 5 shows the major pathways for the many different classes of molecules that can arise from glucosylceramide and galactosylceramide. These are defined by the addition of specific monosaccharide units added in particular sequences. Evidence shows that the early transferases are probably specific for the glycolipids and reside in the early part of the Golgi. In contrast, many of the later-acting ones are probably shared with other types of oligosaccharides and are probably present in the late regions of the Golgi. Thus, as in the case of the *N*-GlcNAc-linked and *O*-GalNAc-linked oligosaccharides, a variety of common outer chains can be mixed and matched among the different classes to generate further diversity. The glycosphingolipids have the greatest variety of outer chains. The more unusual ones are not presented in Figure 5 owing to lack of space. The levels of the different transferases control what kinds of glycosphingolipids are made by a particular cell, but what controls the expression of the enzymes is not known. In contrast to the situation with the *N*-linked oligosaccharide pathway, only one of the glycosphingolipid-specific transferases has been cloned to date.

2.4.3. Termination

As in the case of the *N*- and *O*-linked sugar chains, termination occurs when specific terminal sugars (e.g., sialic acid or blood group sugars) are added. However, details of this process are not clearly understood.

2.4.4. Inhibitors and Mutants

To date, only one specific inhibitor of glycosphingolipid biosynthesis is known. PDMP was specifically synthesized as an inhibitor of glucosylceramide synthase, the enzyme that initiates the production of most major classes of glycosphingolipids. This compound has proven very useful in biosynthetic studies. However, functional studies are somewhat limited by the fact that it can also affect the related pathways for biosynthesis and turnover of sphingomyelin and sphingosine. Most of the available tissue culture mutants that affect glycosphingolipid biosynthesis are pleiotropic ones affecting multiple pathways (see Table III).

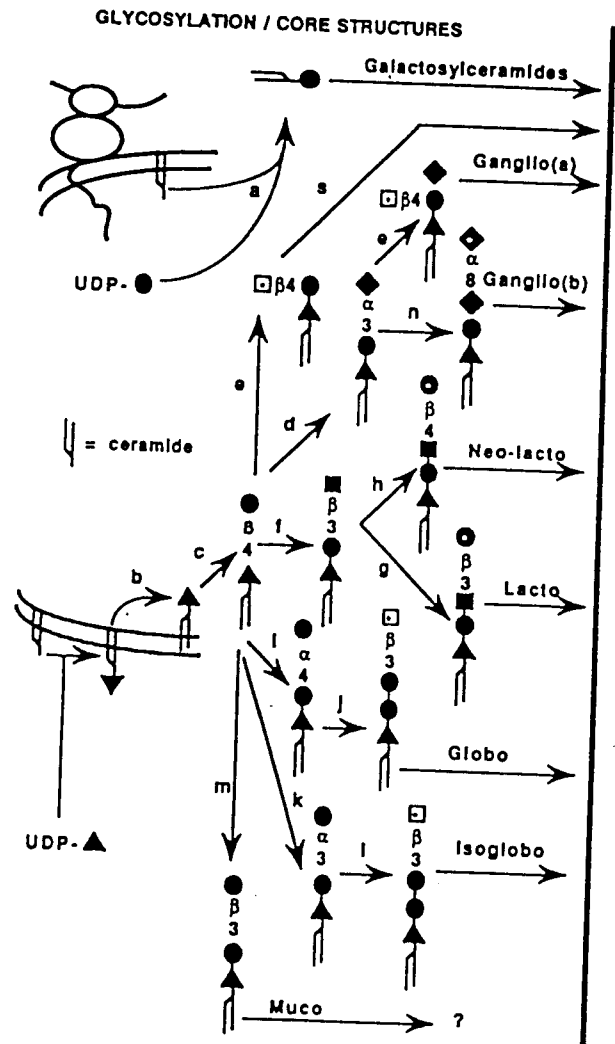
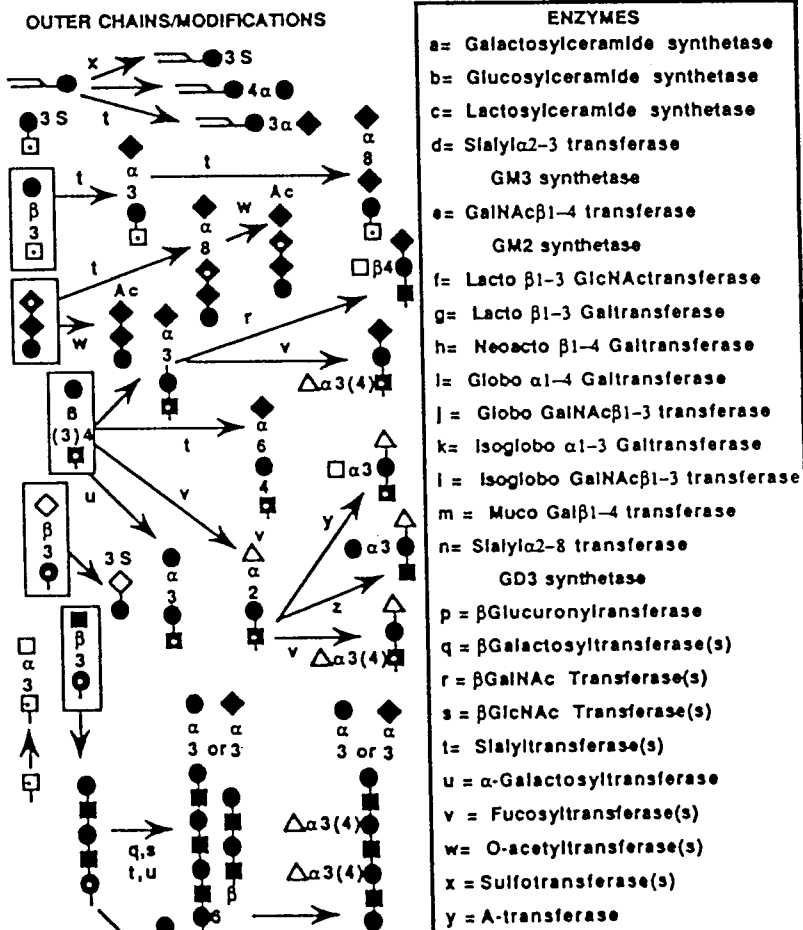
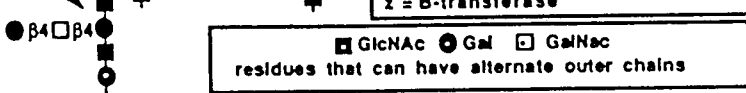


FIGURE 5. The ceramide-linked oligosaccharide

2.4.5. Turnover and Fate

The turnover of glycosphingolipids appears to be rather slow in most cells. There is some evidence that removal and replacement of terminal units can occur, similar to the *N*- and *O*-linked chains. Terminal degradation takes place in the lysosomes and involves not only hydrolases, but also specific "activator" proteins. Glycolipid transfer proteins are also well known and widespread, but their function within the cell is not well understood.





pathway. See Figure 2 for symbol key.

2.5. *O*-Xylose-Linked Chains

The transfer of β -xylose to appropriately disposed serine residues initiates these chains on a limited number of "core proteins." All of these glycans contain a shared linkage region consisting of $\text{GlcUA}\beta 1-3\text{Gal}\beta 1-3\text{Gal}\beta 1-4\text{Xyl-O-Ser}$ (see Figure 6). These cores are extended with alternating disaccharide units of either GlcUA-GlcNAc (heparin and heparan sulfate) or GlcUA-GalNAc (chondroitin sulfate and dermatan sulfate). Many of these glycosaminoglycan (GAG) chains are,

in turn, extensively modified as described below. The finished product, a core protein with one or more GAG chains, is usually called a proteoglycan. Many such distinct core proteins have been cloned, and they can be transmembrane, GPI-anchored (see below), or secreted (Bhavanandan and Davidson, 1992; Kjellén and Lindahl, 1991; Ruoslahti and Yamaguchi, 1991; Yanagishita and Hascall, 1992).

2.5.1. Initiation

Xyl is transferred from UDP-Xyl to proteins that present an available Ser-Gly-X-Gly sequence. This is usually preceded by one or two acidic amino acid residues, but is not a universal signal for addition of Xyl residues. In some proteins, the Ser-Gly dipeptide occurs as a series of contiguous repeats, whereas in others it occurs singly. In the final analysis, Xyl addition is probably determined by a general three-dimensional peptide structure rather than a specific and limited sequence. Some core proteins are "part-time proteoglycans" in which only a portion of the available peptide chains are modified by GAG chains.

2.5.2. Elongation

This can be considered to occur as three distinct groups of reactions.

2.5.2a. Core Synthesis. The GlcUA β 1-3Gal β 1-3-Gal β 1-4Xyl-*O*-Ser core is added sequentially in the early part of the Golgi apparatus by a series of different sugar transferases. In some cases, this is accompanied by phosphorylation of the 2-position of Xyl and sulfation of the 4- or 6-positions of the innermost Gal residue. Both the control and the function of these additional modifications are not understood.

2.5.2b. Chain Extension. The protein itself probably determines whether GlcNAc or GalNAc is added to the GlcUA at the nonreducing end of the glycan core (Figure 6). This decision is crucial because it determines whether the chain will be a heparin/heparan sulfate type or a chondroitin/dermatan sulfate chain. Thereafter, the chains are polymerized by the sequential addition of each monosaccharide from its nucleotide donor. Some core proteins can alternately express chondroitin sulfate or heparin/heparan sulfate as the glycan, depending upon the cell in which they are synthesized.

2.5.2c. Chain Modification. The individual chains are modified during their elongation in a variety of ways shown in Figure 6. Not all of the residues are modified identically, and the control of this modification is not well understood. In the case of heparin and heparan sulfate, selected GlcNAc residues are first de-*N*-acetylated and then *N*-sulfated. At this point, some of the GlcUA can be epimerized to IdUA; *O*-sulfate esters can be added to GlcNAc, GlcNSO₄, or IdUA residues. In the case of the chondroitin sulfate chains, little GlcUA is epimerized, and GalNAc can be sulfated on the 4- and/or 6-positions. In der-

matan sulfate, a portion of GlcUA can be epimerized to IdUA. All of these highly ordered and interdependent modifications happen as the proteins are rapidly moving through the late part of the Golgi apparatus.

2.5.3. Termination

Very little is known about the factors that control the size of the chain or how it is terminated.

2.5.4. Turnover and Fate

In some cell types, proteoglycans are segregated into secretory vesicles, where the glycosaminoglycan chains are depolymerized prior to controlled secretion. Cell surface or secreted proteoglycans can be taken up and degraded either by the same cell that made them, or by another cell. After uptake, some GAG chains are believed to be directed to distinct cellular compartments such as the nucleus. However, all GAG chains must eventually be broken down to their component parts in the lysosomes. Considerable evidence shows that the breakdown of GAG chains in the lysosomes occurs in a highly ordered sequence.

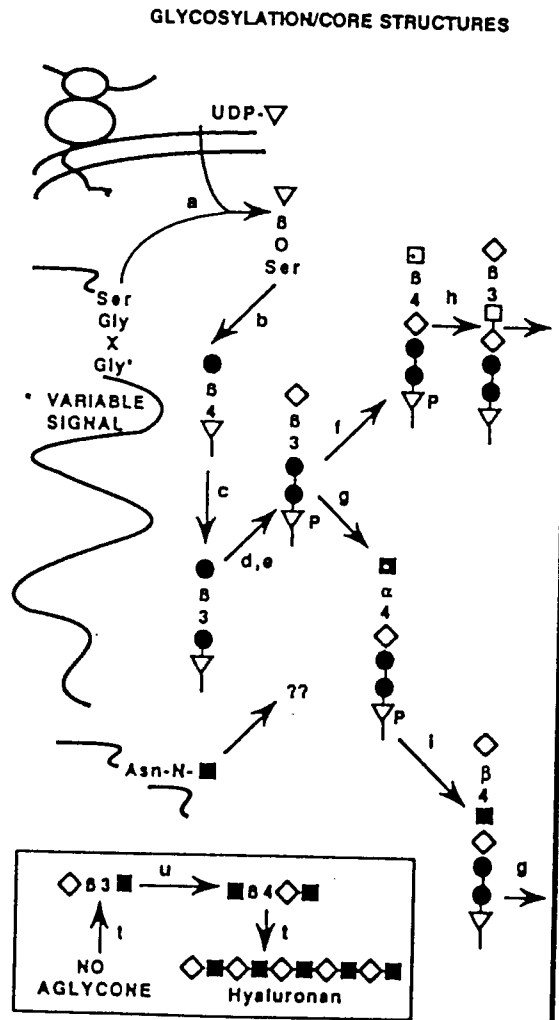
2.5.5. Inhibitors and Mutants

Proteoglycan synthesis can be inhibited and GAG chain synthesis greatly stimulated by feeding cells a variety of different β -xylosides. Paranitrophenyl, 4-methyl-umbelliferyl and even methyl-xyloside derivatives at submillimolar concentrations can penetrate the cells and the Golgi, where they serve as alternative acceptors for the addition of the first Gal residue. Because the xylosides are in such high concentration compared to the core proteins, they are very effective competitors. Further elongation of these chains occurs, but is heavily biased toward adding chondroitin/dermatan sulfate chains rather than heparan sulfate chains. These chains are generally smaller than those found on natural core proteins. Recently, some evidence shows that the structure of the aglycone can determine which type of GAG chain is added to the artificial core.

Several mutant CHO cell lines have been isolated that either cannot make a carbohydrate core region or do not elongate the chains. The core-deficient mutants lack xylosyl or galactosyltransferase (Figure 6). Others lack one of the chain polymerizing enzymes, or a specific sulfotransferase.

2.6. Glycophospholipid Anchors

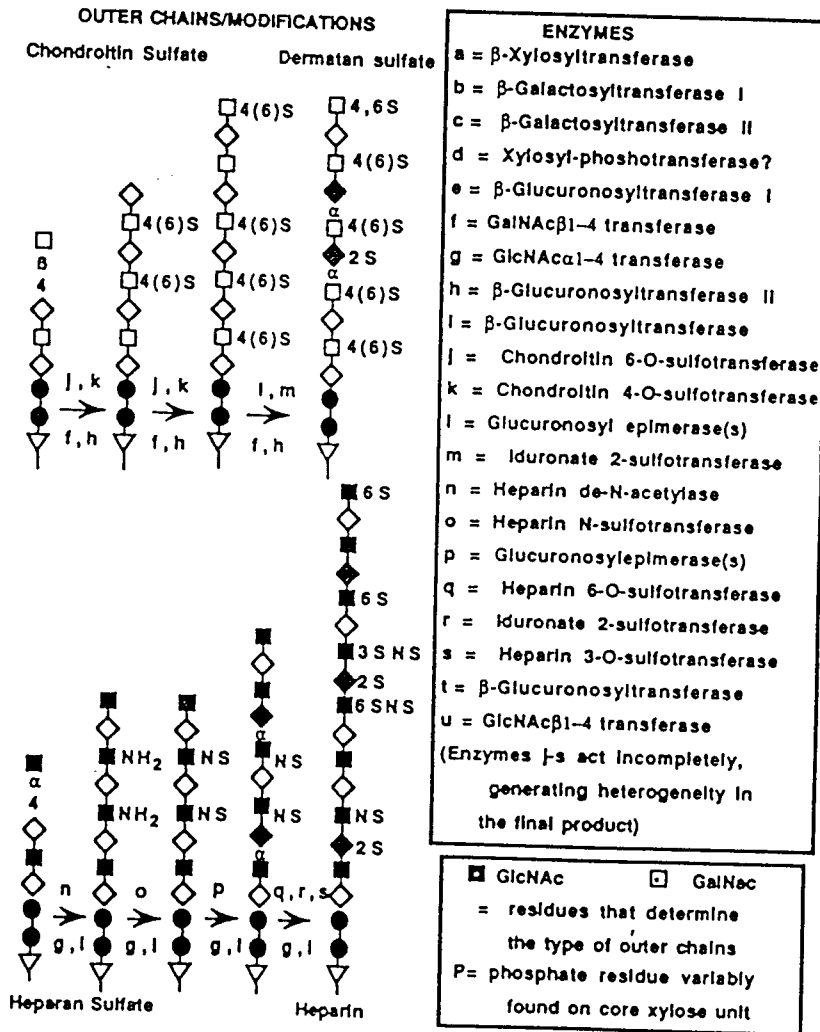
These structures replace the carboxyl-terminal transmembrane portions of certain proteins, and thus serve as alternate ways to anchor such proteins to cell surfaces. Some proteins can exist either with transmembrane regions or with glycophospholipid (GPI) anchors. In some cases alternative splicing of mRNA



can determine the outcome. The common core structure of these anchors can be elongated or modified in several ways (see Figure 7). These core structures can be cleaved by specific microbial and mammalian phospholipases (Doering *et al.*, 1990; Ferguson, 1992).

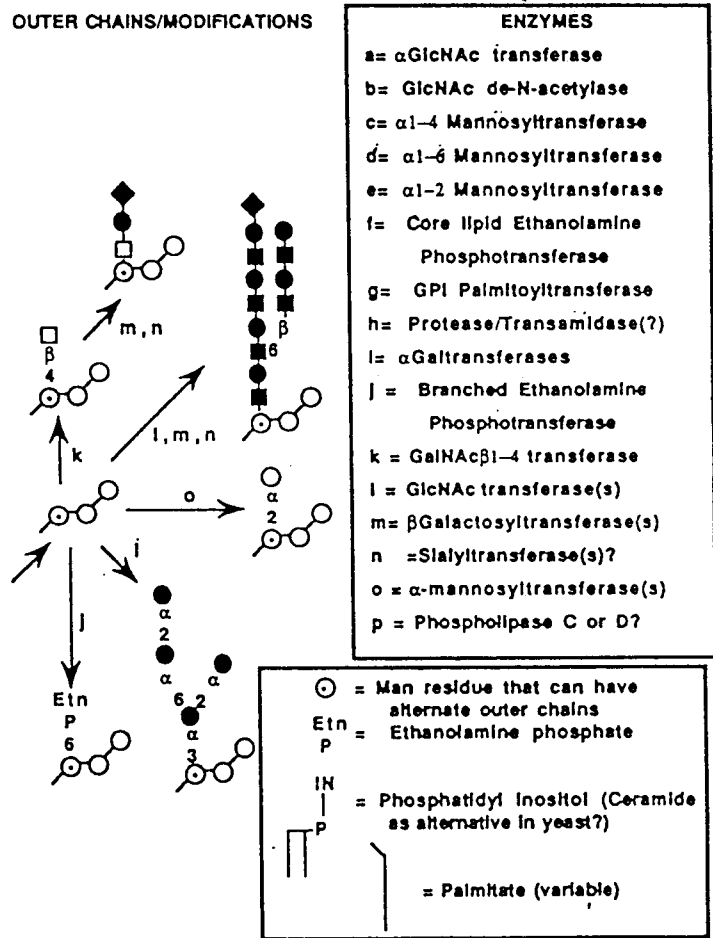
2.6.1. Initiation

The core anchor structure has several unique features, including the presence of ethanolamine and a nonacetylated glucosamine. It is both performed and



pathway. See Figure 2 for symbol key.

used in the ER to replace the transmembrane region of certain proteins in a transamidation reaction. The acceptor sequence for this reaction is found at the carboxyl terminus of newly synthesized proteins and appears to include specific residues about 20 amino acids from the membrane. In some cells, a fatty acyl chain can also be added to the core inositol group, rendering the anchor resistant to phospholipases. Also, additional ethanolamine phosphate residues can be added to the core mannose residues in some cell types (not shown in Figure 7).



pathway. See Figure 2 for symbol key.

2.6.3. Termination

Apart from the side chain, no specific processing of the core structure is known. In some cases the fatty acyl chain added to the inositol ring can be removed.

2.6.4. Inhibitors and Mutants

A variety of mutants affect several steps in the assembly of the GPI anchors. These include both tissue-culture cell lines and cells from patients with an ac-

Table IV
Less Common Types of Carbohydrate-Protein Linkages

Linkage	Found in
GlcNAc- <i>O</i> -Serine	Cytosolic glycoproteins, transcription factors, cytokeratins, RNA polymerase, cytoskeletal proteins
GLC-P-Man- <i>O</i> -Serine	Parafuscin, a cytosolic glycoprotein.
Fuc- <i>O</i> -Serine	Tissue plasminogen activator
Xyl-Xyl-Glc- <i>O</i> -Serine	Clotting factors
GlcNAc-1-P-Serine	Proteinase in <i>Dictyostelium</i>
Glc- <i>O</i> -Tyrosine	Glycogen core protein
Gal-Glc- <i>O</i> -Hydroxylysine	Collagen
Man- <i>O</i> -Ceramide	Lower invertebrate glycosphingolipids

quired disorder of blood cells called paroxysmal nocturnal hemoglobinuria (PNH). Because dolichol-phosphoryl mannose is a donor in the synthesis of the anchor, some of the mutants affect the biosynthesis of both *N*-linked oligosaccharides and GPI anchors.

2.6.5. Fate and Turnover

Little is known about this matter. It is suggested that, in some cases, phospholipases at the cell surface can release the protein with the oligosaccharide attached to its carboxyl terminus. Also, some controversial evidence indicates that by-products of this pathway may act as second messengers for certain hormones.

2.7. Less Common Types of Oligosaccharide Linkages

The types of glycoconjugates discussed to this point are by far the most common forms encountered, but they are not the only ones. Table IV lists some of the rarer forms of protein-carbohydrate linkages. These types are not as well studied as the commoner ones (except in the case of *O*-linked GlcNAc residues added in the cytoplasm, which are discussed below). It is important to keep in mind that a rarer oligosaccharide is also unique, and may be restricted to a particular protein or class of proteins. As such, it would be theoretically more likely to carry out some unique biological function(s) (Kornfeld, 1990; Troy, 1992).

2.8. Glycoconjugates Frequently Share Identical Outer Sequences

In contrast to the very distinctive features of the core regions of the different types of oligosaccharides, a large number of outer sequences are shared by a number of glycoconjugates (Kobata and Takasaki, 1992; Fukuda, 1992). These

STRUCTURE	N-GlcNAc-Linked	O-GalNAc-Linked	Glc-Cer-Linked	GPI-Linked	Common Name(s)
●β4■β-	+	+	+	+	Type 2 Lactosamine
◆α6●β4■β-	+	+	+		CD75, ?CD76
9(7)(4)A c◆α6●β3■β-	+	+			
◆α3●β4■β-	+	+	+		
◆α8◆α3●β4■β-	+	+			
(◆α8◆) _n α3●β4■β-	+	+			Polysialic Acid (PSA)
●α3●β4■β-	+		+		
△α2●β4■β-	+	+	+		
●β4■β- α3 △	+	+	+		Lewis ^x (SSEA-1) CD15
△α2●β4■β- α3 △	+	+	+		Lewis ^y
◆α3●β4■β- α3 △	+	+	+		Sialyl Lewis ^x
◆α6●β4■β- 6S	+	+			
◆α3●β4■β- 6S	+ ?	+			
8S◆α8◆α3●β4■β-			+		
3S●β4■β-	+				
3S◇β3●β4■β-	+ ?		+		HNK-1 antigen
◆α3●β4■β- β4 □	+	+	+		Sda/CAD antigen
●β4■β3(●β4■β3) _n	+	+	+	+	Poly-lactosamine (i antigen)
●β4■β3●β4■β3●β4■β3 β6 ●β4■β3●	+	+	+	+ ?	Branched Poly-lactosamine (I antigen)
●β4■β3●β4■β- α3 △	+	+	+		Dimeric Lewis ^x
△α2●β4■β3●β4■β- α3 △	+	+	+		Dimeric Lewis ^y
◆α3●β4■β3●β4■β- α3 △	+	+	+		Sialyl dimeric Lewis ^x
◆α3●β4■β3●β4■β- α3 △	+	+	+		VIM-2, CD65
◆α6●β4■β3●β4■β- α3 △	+	+	+		

FIGURE 8. Some outer chain sequences found on different types of glycans. See Figure 2 for symbol key.

STRUCTURE	N-GlcNAc-Linked	O-GalNAc-Linked	Glc-Cer-Linked	GPI-Linked	Common Name(s)
$\blacklozenge \alpha 3(6) \bullet \beta 4 \blacksquare \beta 3 \bullet \beta 4 \blacksquare \beta 3$ 6S 6S 6S	+	+			Keratan Sulfate
$\bullet \beta 3 \blacksquare \beta -$	+	+	+		Type 1 lactosamine
$\blacklozenge \alpha 3 \bullet \beta 3 \blacksquare \beta -$	+	+	+		
$\bullet \alpha 3 \bullet \beta 3 \blacksquare \beta -$	+	+	+		
$\triangle \alpha 2 \bullet \beta 3 \blacksquare \beta -$	+	+	+		
$\bullet \beta 3 \blacksquare \beta -$ $\alpha 4$	+	+	+		Lewis ^a
$\triangle \alpha 2 \bullet \beta 3 \blacksquare \beta -$ $\alpha 4$	+	+	+		Lewis ^b
$\triangle \alpha 2 \bullet \beta 3 \blacksquare \beta -$ $\alpha 3$ $\alpha 4$	+	+	+		Blood Group B
$\triangle \alpha 2 \bullet \beta 3 \blacksquare \beta -$ $\alpha 3$ $\alpha 4$ \square	+	+	+		Blood Group A
$\blacklozenge \alpha 3 \bullet \beta 3 \blacksquare \beta -$ $\alpha 4$	+	+	+		Sialyl Lewis ^a
$\blacklozenge \alpha 3 \bullet \beta 3 \blacksquare \beta -$ $\alpha 6$	+		+		
$\blacklozenge \alpha 3 \bullet \beta 3 \blacksquare \beta -$ $\alpha 6$ $\alpha 4$			+		Disialyl Lewis ^a
$\square \beta 4 \blacksquare \beta -$	+				
$4S \square \beta 4 \blacksquare \beta -$	+				
$\blacklozenge \alpha 6 \square \beta 4 \blacksquare \beta -$	+				
$\bullet \beta 3 \square$		+	+	+	PNA receptor
$\blacklozenge \alpha 3 \bullet \beta 3 \square$		+	+	+	
$\blacklozenge \alpha 8 \blacklozenge \alpha 3 \bullet \beta 3 \square$			+		
$\bullet \alpha 3 \bullet \alpha 3$			+		
$\square \beta 3 \bullet \alpha -$			+		
$\square \alpha 3 \square \beta 3 \bullet \alpha -$			+		Forsman Antigen
$\circ \alpha 2 \circ \alpha -$	+			+	

FIGURE 8. (continued)

include sialic acids, polylactosamines, blood group sequences, etc. (see Figures 3, 5, and 8). These outer chains are probably added to the different acceptors (O- or N-linked chains, ceramide-linked chains, etc.) by overlapping and, in some cases, probably identical transferases. How this is controlled from one glycoconjugate to another is not known, but it underscores the importance of the initiation

Table V
Examples of Modifications of the Oligosaccharide Chains
on Animal Glycoconjugates

Type	Modified sugar	Found in
Phosphate	Man-6-P	Lysosomal enzymes, TGF- β precursor
	GlcNAc-6-P	Milk oligosaccharides
	Xyl-2-P	Core xylose of glycosaminoglycans
	Inositol-1-P	Glycophospholipid anchors
	Sialic acid-9-P	Biosynthetic intermediate (free)
	GlcNAc-1-P-6 Man	Precursor of Man-6-P (see above)
	Glc-1-P-6-Man	Cytosolic parafuscin or phosphoglucomutase
	GlcNAc-1-P-Ser	Dictyostelium proteins
	GalNAc-1-P?	<i>N</i> -cadherin and ? other proteins
	Man-1-P-6-Man	Yeast mannans
Sulfate	Gal-3-S	Sulfatide (ceramide-linked) thyroglobulin (<i>N</i> -GlcNAc-linked)
	Gal-6-S	Keratan sulfate
	Man-6-S	Dictyostelium lysosomal enzymes
	Man-3(4)-S	Ovalbumin
	GlcNAc-6-S	Heparin/heparan sulfate, <i>N</i> -GlcNAc-linked chains (endothelium), keratan sulfate
	GlcNS	Heparin/haparan sulfate
	GlcNS-3-S	Heparin/heparan sulfate
	GlcNS-3,6-S	Heparin/haparan sulfate
	GalNAc-4-S	Chondroitin sulfate
	GalNAc-6-S	Chondroitin sulfate
	GlcA-3-S	Heparin sulfate
	Id-2-S	Heparin sulfate
	GlcUA-3-S	Glycosphingolipids, ? <i>N</i> -GlcNAc-linked oligosaccharides
	Sia-8-S	Glycosphingolipids: echinoderm, bovine
<i>O</i> -Acetyl	Sialic acid-4,7,8,9	<i>O</i> - <i>N</i> - and Ceramide-linked oligosaccharides
<i>O</i> -Methyl	Sialic acid-8	Echinoderm glycosphingolipids
<i>O</i> -Acyl	Inositol	Palmitoylation on GPI anchors
	Galactose	Acylated Gal-Ceramide in brain
	Sialic acid	<i>O</i> -lactyl groups on mucins and on free sialic acids
Ethanolamine-P	Mannose	GPI anchors

of the sugar chains. It should not be surprising that similar outer structures could be used to accomplish different functions when presented in different contexts. Tissue-specific expression of different genes whose products catalyze the same reaction but with differing substrate requirements may also be used to generate "context-specific" glycosylation.

2.9. Modifications of the Sugar Units in Oligosaccharide Chains

The different types of sugar chains described here are subject to further modification by the addition of a variety of different types of substituents (e.g., phosphate, sulfate, acetate, etc.). A list of known modifications is given in Table V. Because the addition of such substituents imparts further diversity to the common core oligosaccharide structures, it is reasonable to speculate that they may impart more specific functions to the sugar chains. In at least some cases (e.g., mannose-6-phosphate on lysosomal enzymes and specific types of sulfation in heparin), this is known to be the case (Kornfeld, 1990; Cumming, 1991).

3. NUCLEAR AND CYTOPLASMIC GLYCOSYLATION

This chapter has focused upon glycosylation that occurs in the ER-Golgi-plasmalemma pathway. Until a few years ago it was thought that glycosylation did not occur on molecules on the opposite side of the membrane (i.e., the cytosol), which is contiguous with the interior of the nucleus. However, it is now clear that the addition of single *O*-GlcNAc residues is very common on cytosolic and nuclear proteins (Hart *et al.*, 1989). Some proteins have many such units, sometimes closely spaced. Clues from many investigators suggest that this may only be the proverbial tip of the iceberg, and that many other types of nuclear and cytoplasmic glycosylation may occur. This topic deserves more attention than can be accorded it here.

ACKNOWLEDGMENTS. This work was supported by USPHS grants GM32373, CA38701, and GM32485, and a VA Merit Review Award. H. F. is an established investigator of the American Heart Association. We thank Michael Bevilacqua, Christoph Binkert, Daniel Donoghue, Richard Dutton, Michiko Fukuda, Chris Glass, Robert Linhardt, Elaine Muchmore and Harry Schachter for their helpful comments and suggestions.

4. REFERENCES

- Bhavanandan, V. P., and Davidson, E. A., 1992, Proteoglycans: Structure, synthesis, function, in *Glycoconjugates: Composition, Structure and Function* (H. J. Allen and E. C. Kisailus, eds.), pp. 167-202. M. Dekker, Inc., New York.
- Carraway, K. L., and Hull, S. R., 1991, Cell surface mucin-type glycoproteins and mucin-like domains, *Glycobiology* 1:131-138.
- Cumming, D. A., 1991, Glycosylation of recombinant protein therapeutics: Control and functional implications, *Glycobiology* 1:115-130.
- Cummings, R. D., Merkle, R. K., and Sultz, N. L., 1989, Separation and analysis of glycoprotein oligosaccharides, *Methods Cell Biol.* 32:141-183.

- Cummings, R. D., 1992, Synthesis of asparagine-linked oligosaccharides: Pathways, genetics, and metabolic regulation, in *Glycoconjugates: Composition, Structure and Function* (H. J. Allen and E. C. Kisailus, eds), pp. 333-360, M. Dekker, Inc., New York.
- Doering, T. L., Masterson, W. J., Hart, G. W., and Englund, P.T., 1990, Biosynthesis of glycosyl phosphatidylinositol membrane anchors, *J. Biol. Chem.* 265:611-614.
- Elbein, A. D., 1991, Glycosidase inhibitors: Inhibitors of N-linked oligosaccharide processing, *FASEB J.* 5:3055-3063.
- Farquhar, M. G., 1991, Protein traffic through the Golgi complex, in *Intracellular Trafficking of Proteins* (C. J. Steer and J. Hanover, eds.), pp. 431-471, Cambridge University Press, New York.
- Ferguson, M. A. J., 1992, Lipid anchors on membrane proteins, *Curr. Opin. Struct. Biol.* 1:522-529.
- Fukuda, M., 1992, Cell surface carbohydrates in hematopoietic cell differentiation and malignancy, in *Cell Surface Carbohydrates and Cell Development* (M. Fukuda, ed.), pp. 127-159, CRC Press, Boca Raton, Florida.
- Hakomori, S., 1986, Tumor associated glycolipid antigens, their metabolism and organization, *Chem. Phys. Lipids* 42:209-233.
- Hakomori, S., 1990, Bifunctional role of glycosphingolipids: Modulators for transmembrane signaling and mediators for cellular interactions, *J. Biol. Chem.* 265:18713-18716.
- Hart, G. W., Haltiwanger, R. S., Holt, G. D., and Kelly, W. G., 1989, Glycosylation in the nucleus and cytoplasm, *Annu. Rev. Biochem.* 58:841-874.
- Hirschberg, C. B., and Snider, M. D., 1987, Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus, *Annu. Rev. Biochem.* 56:63-87.
- Kisailus, E. C., and Allen, H. J., 1992, Nomenclature of monosaccharides, oligosaccharides, and glycoconjugates, in *Glycoconjugates: Composition, Structure and Function* (H. J. Allen and E. C. Kisailus, eds.), pp. 13-32, M. Dekker, Inc., New York.
- Kjellén, L., and Lindahl, U., 1991, Proteoglycans: Structures and interactions, *Annu. Rev. Biochem.* 60:443-475.
- Kobata, A., and Takasaki, S., 1992, Structure and biosynthesis of cell surface carbohydrates, in *Cell Surface Carbohydrates and Cell Development* (M. Fukuda, ed.), pp. 1-24, CRC Press, Boca Raton, Florida.
- Kornfeld, R., and Kornfeld, S., 1985, Assembly of asparagine-linked oligosaccharides, *Annu. Rev. Biochem.* 54:631-664.
- Kornfeld, S., 1990, Lysosomal enzyme targeting, *Biochem. Soc. Trans.* 18:367-374.
- Kundu, S. K., 1992, Glycolipids: Structure, synthesis, functions, in *Glycoconjugates: Composition, Structure and Function* (H. J. Allen and E. C. Kisailus, eds.), pp. 203-262, M. Dekker, Inc., New York.
- Mellman, I., and Simons, K., 1992, The Golgi complex: *In vitro* veritas? *Cell* 68:829-840.
- Paulson, J. C., and Colley, K. J., 1989, Glycosyltransferases: Structure, localization, and control of cell type-specific glycosylation, *J. Biol. Chem.* 264:17615-17618.
- Rademacher, T. W., Parekh, R. B., and Dwek, R. A., 1988, Glycobiology, *Annu. Rev. Biochem.* 57:785-838.
- Rothman, J. E., and Orci, L., 1992, Molecular dissection of the secretory pathway, *Nature* 355:409-415.
- Ruoslahti, E., and Yamaguchi, Y., 1991, Proteoglycans as modulators of growth factor activities, *Cell* 64:867-869.
- Schachter, H., 1991, The "yellow brick road" to branched complex N-glycans, *Glycobiology* 1:453-462.
- Schachter, H., and Brockhausen, I., 1992, The biosynthesis of serine (threonine) N-acetylglucosamine-linked carbohydrate moieties, in *Glycoconjugates: Composition, Structure and Function* (H. J. Allen and E. C. Kisailus, eds.), pp. 263-332, M. Dekker, Inc., New York.

- Schauer, R., 1991, Biosynthesis and function of *N*- and *O*-substituted sialic acids. *Glycobiology* 1:449-452.
- Shayman, J. A., and Radin, N. S., 1991, Structure and function of renal glycosphingolipids. *Am. J. Physiol. Renal, Fluid Electrolyte Physiol.* 260:F291-F302.
- Stanley, P., 1992, Glycosylation engineering. *Glycobiology* 2:99-107.
- Stults, C. L. M., Sweeley, C. C., and Macher, B. A., 1989, Glycosphingolipids: Structure, biological source, and properties. *Methods Enzymol.* 179:167-214.
- Troy, F. A., II, 1992, Polysialylation: From bacteria to brains. *Glycobiology* 2:5-23.
- VanEchten, G., and Sandhoff, K., 1993, Ganglioside metabolism: Enzymology, topology, and regulation. *J. Biol. Chem.* 268:5341-5344.
- Varki, A., 1991, Radioactive tracer techniques in the sequencing of glycoprotein oligosaccharides. *FASEB J.* 5:226-235.
- Varki, A., 1992a, Diversity in the sialic acids. *Glycobiology* 2:25-40.
- Varki, A., 1992b, Role of oligosaccharides in the intracellular and intercellular trafficking of mammalian glycoproteins, in *Cell Surface Carbohydrates and Cell Development* (M. Fukuda, ed.), pp. 25-69. CRC Press, Boca Raton, Florida.
- Varki, A., 1993, Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology* 3:97-130.
- Winchester, B., and Fleet, G. W. J., Amino-sugar glycosidase inhibitors: Versatile tools for glycobologists. *Glycobiology* 2:199-210.
- Yanagishita, M., and Hascall, V. C., 1992, Cell surface heparan sulfate proteoglycans. *J. Biol. Chem.* 267:9451-9454.