56 Glycoproteins

Robert K Murray, MD, PhD

INTRODUCTION

Glycoproteins are proteins that have oliosaccharide (glycan) chains covalently attached to their polypeptide backbones. Glycoproteins are one class of **glycoconjugate** or **complex carbohydrates.** These are **equivalent** terms used to denote molecules containing one or more carbohydrate chains covalently linked to protein (to form glycoproteins or proteoglycans) or lipid (to form glycolipids). **Proteoglycans** are discussed in Chapter 57 and **glycolipids** in Chapter 16.

BIOMEDICAL IMPORTANCE

Almost all the **plasma proteins** of humans, except albumin are glycoproteins. Many **proteins of cellular membranes** (Chapter 43) contain substantial amounts of carbohydrate. **A** number of the **blood group substances** are glycoproteins, whereas others are glycosphingolipids. Certain **hormones** (eg, choronic gonadotropin) are glycoproteins. **Cancer** is increasingly recognised as a disorder resulting from abnormal genetic regulation (Chapter 62). The major problem in cancer is **metastasis,** the phenomenon whereby cancer cells leave their tissue of origin (eg, the breast), migrate through the blood stream to some distant site in the body (eg, the brain), and grow there in a completely unregulated manner, with cata strophic results for the affected individual. Many cancer researchers think that alterations in the structures of glycoproteins and other glyconjugates on the surfaces of **cancer** cells are important in the phenomenon of metastasis.

GLYCOPROTEINS OCCUR WIDELY AND PERFORM NUMEROUS FUNCTIONS

Glycoproteins occur in most organisms, from bacteria to humans. Many animal viruses also contain glycoproteins, some of which have been much investigated in part because they are very suitable, for biosynthetic studies. Numerous proteins with diverse fuctions are glycoproteins (Table 56- 1);

Table 56-1. Some fuctions served by glycproteins

Function	Glycoproteins
Lubricant ad protective agent	Mucins
Transport molecule	Transferrin, ceruloplasm
Immunologic molecule	Immunoglobulins, histocompatibility antigens
Hormone	Chorionic gonadotropin, thyroid-stimulating hormone (TSH)
Enzyme	Various, eg, alkaline phosphatase
Cell attachment-recognition site	Various proteins involved in cell-cell (eg, sperm-oocyte), virus-cell, bacterium cell, and hormone-cell interactions.
Antifreeze	Certain plasma proteins of cold water fish.
Interact with specific carbohydrates	Some lectins

Their carbohydrate content ranges from 1% to over 85% by weight.

Many studies have been performed to define the precise roles oligosaccharide chains play in the functions of glycoproteins. Table 56-2 summarises results from such studies; some of the functions listed are firmly established, and others are still under investigation.

OLIGOSACCHARIDE CHAINS ENCODE BIOLOGIC INFORMATION

An enormous number of glycosidic linksages can be generated between sugars. For example, three differentt hexoses may he linked to each other to form over one thousand different trisaccharides. The conformations of the sugars in oligosaccharide chains vary depending on their linkages and proximity to other molecules with which the oligosaccharides may interact. A widely held belief is that oligosaccharide chains encode considerable **biologic information** and that this depends upon the constituent sugars, their sequences, and their conformations. For instance, mannose 6-

Table 56-2. Some functions of the oligosaccharide chains of glycoproteins. 1

- Modulate physiochemical properties, e.g., solubility, viscosity, charge, and denaturation
- Protect against proteolysis, from inside and outside of cell
- Affect proteolitic processing of precursor proteins to smaller products.
- Are involved in biologic activity, e.g., of human chorioic gonadotrophin (hCG)
- Affect insertion into membranes, intracellular migration, sorting and secretion
- Affect embryonic development and differentiation
- May affect sites of metastases selected by cancer cells.

I.Adapted from Schachter FL Biosynthetic controls that deter-rmine the branching and heterogeneity of protein-bound oligosaccharides. Biechem Cell Biol 1936;64:163.

phosphate residues **target** newly synthesised lysosomal enzymes to that organaelle (see below).

TECHNIQUES ARE AVAILAIBLE FOR DETECTION, PURIFICATION, AND STRUCTURAL ANALYSIS OF GLYCOPROTEINS

A variety of methods that are used in the detection, purification and structural analysis of glycoproteins are listed in Table 56-3. The conventional methods used to purify proteins and enzymes (Chapter 8) are also applicable to the purification of glycoproteins. Once a glycoprotein has been **purified,** the use of **mass spectrometry"** and **high-resolution NMR spectroscopy** can often identify the structures of the glycan chains present in a glycoprotein. Analysis of glycoproteins can be complicated by the fact that they often exist as glycoforms; these are proteins with identical amino acid sequence but different oligosaccharide compositions. Although linkage details are not stressed in this particular chapter, it is critical to appreciate that the precise natures of the **linkages** between the sugars of glycoproteins are of fundamental importance in determining the structures and functions of these molecules.

EIGHT SUGARS PREDOMINATE IN HUMAN GLYCOPROTEINS

About 200 monosaccharides are found in nature; however, only eight are commonly found in the oligosaccaride chains of glycoproteins (Table 56- 4). Most of these sugars were described in Chapter 15. N-Acetylneuraminic acid (NcuAc) is found at the terminals of oligosaccharide chains, usually attached to subterminal galactose (Gal) or N-acetyl-galactosamine (GaINAe) residues. The other sugars listed are generally found in more internal positions. **Sulfate** is often found in glycoproteins, usually attached to Gal GaINAc, or GlcNAc.

NUCLEOTIDE SUGARS ACT AS SUGAR DONORS IN MANY BIOSYNTHE'UC REACTIONS

The first nucleotide sugar to be reported was uridine diphosphate glucose (UDPG1c); its structure is shown in Figure 20-2. The common nucleotide sugars involved in the biosynthesis of glycoproteins are listed in Table 56-4; the reasons some contain UDP

Table 56-3. Some important methods used to study glycoproteins

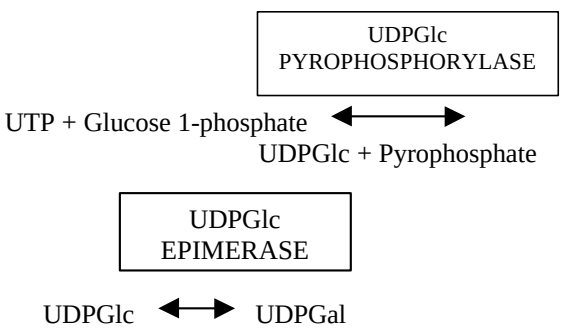
Method	Use
Periodic acid shift reagent	Detects glycoproteins as pink bands after electrophoretic separation.
Incubation of cultured cells with a radioactive sugar	Leads to detection of glycoproteins as radioactive bands after electrophoretic separation.
Treatment with appropriate glycosidase or phosphilipase.	Resultant shifts in electrophoretic migration help distinguish among proteins with N-glycan, O-glycan, or GPI linkages and also between high mannose and complex N-glycans.
Sepharose-lectin column chromatography	To purify glycoproteins or glycopeptides that bind the particular lectin used.
Compositional analysis following acid hydrolysis	Identifies sugars that the glycoprotein contains and their stoichiometry.
Mass Spectrometry	Provides information on molecular mass, composition, sequence and sometimes branching of a glycan chain.
NMR Spectroscopy	To identify specific sugars, their sequence, linkages and the aoremic nature of glycosidic linkages.
Methylation (linkage) analysis	To determine linkages between sugars.
Amino Acid or cDNA sequencing	Determination of Amino Acid sequence.

Table 56-4. The principal sugars found in human glycoproteins. Their structures are illustrated in Chapter 15.

Sugar	Type	Abbreviation	Nucleotide Sugar	Comments
Galactose			UDPGal	Often found subterminal to NeuAc in N-linked glycoproteins. Also found in core trisaccharide of proteoglycans.
Glucose			UDPGlc	Present during the glycosynthesis of glycoproteins but not usually present in mature glycoproteins.
Mannose	Hexose	Man	GDP-Man	Common sugar in N-linked glycoproteins.
N-Acetyleuraminic Acid	Sialic Acid (nine C atoms)	NeuAc	CMP-NeuAc	Often the terminal sugar in both N-liked and O-linked glycoproteins. Other types of Sialic acid are also found, but NeuAc is the major species found in humans.
Fucose	Deoxyhexose	Fuc	GDP-Fuc	May be external in both N-liked and O-linked glycoproteins or linked to the GlcNAc residue attached to the Asn in N-linked species.
N-Acetylgalactosamine	Aminohexose	GalNAc	UDP-GalNAc	Present in both N- and O-linked glycoproteins.
N-Acetylglucosamine	Aminohexose	GlcNAc	UDP-GlcNAc	The sugar attached to the polypeptide chain via Asn in N-linked glycoproteins; also found at other sites in the oligosaccharides of these proteins.
Xylose	Pentose	Xyl	UDP-Xyl	Xyl is attached to the OH of Ser in many proteoglycans. Xyl in turn is attached to two Gal residues forming a link trisaccharide

and others guanosine diphosphate (GDP) or cytidine monophosphate (CMP) are obscure. Many, but not all, of the glycosylation reactions involved in the biosynthesis of glycoproteins utilize these compounds (see below). The anhydro nature of the linkage between the phosphate group and the sugars is of the high-energy, high-group-transfer-potential type (Chapter 12). The sugars of these compounds are thus "activated" and can be transferred to suitable acceptors provided appropriate transferases are available.

The nucleotide sugars are formed in the cytosol, generally from reactions involving the corresponding nucleoside triphosphate. Formation of uridine diphosphate galactose (UDPGal) requires the follow-ing two reactions:



Because many glycosalation reactions occur within the lumen of the Golgi apparatus, carrier systems (permeases, transporters) are necessary to transport nucleotide sugars across the Golgi membrane.

Systems transporting UDPGal, GDP-Man, and CMP-NeuAc into the cisternae of the Golgi apparatus have been described. They are antiport systems; ie, the influx of one molecule of nucleotide sugar is balanced by the efflux of one molecule of the corresponding nucleotide (eg, UMP, GMP, or CMP) formed from the nucleotide sugars. This mechanism ensures an adequate concentration of each nucleotide sugar inside the Golgi. UMP is formed from UDPGal in the above process as follows: