*Chapter 5*

# Biochemistry and Oncology of Sialoglycoproteins

Veer P. Bhavanandan and Kiyoshi Furukawa

## 1. GENERAL INTRODUCTION

Sialoglycoproteins are defined as proteins containing one or more sialyl oligosaccharides covalently bound via their reducing end to the polypeptide chain. Most, but not all, of the glycosylated proteins in animals contain sialic acid and therefore qualify as sialoglycoproteins. Some examples of glycoproteins lacking sialic acid are ovalbumin, ribonuclease, and antifreeze glycoprotein. Sialoglycoproteins are ubiquitous in animals, either as components of cellular membranes of extracellular fluids such as serum, spinal fluid, saliva, respiratory mucous, gastric juice, sweat, and semen. The intrinsic membrane glycoproteins of the cell have their carbohydrate moieties usually projecting into the cytoplasm, or in the case of the plasma membrane, the exterior of the cell. The cell membrane sialoglycoproteins and sialoglycolipids (gangliosides) enrich the cell surface in sialyl residues which are important determinants in the social behavior of the cell.

---

**Veer P. Bhavanandan** Department of Biochemistry and Molecular Biology, The Milton S. Hershey Medical Center, Pennsylvania State University, Hershey, Pennsylvania 17033. **Kiyoshi Furukawa** Department of Biochemistry, Institute of Medical Science, University of Tokyo, Tokyo 108, Japan; *present address:* Department of Biosignal Research, Tokyo Metropolitan Institute of Gerontology, Tokyo 173, Japan.

*Biology of the Sialic Acids,* edited by Abraham Rosenberg. Plenum Press, New York, 1995.

## 2. CLASSIFICATION AND NOMENCLATURE

Sialoglycoproteins tend to be classified according to the nature of the predominant oligosaccharide–protein linkages. Those in which the majority of the oligosaccharides are linked, via terminal GlcNAc, to the amide group of asparagine are referred to as Asn- or *N*-linked glycoproteins, *N*-glycosylated proteins, *N*-glycans, or serum-type glycoproteins since the majority of the serum components are sialoglycoproteins of this class. Those in which the oligosaccharides are linked via GalNAc to the hydroxyl group of serine and threonine are referred to as Ser/Thr- or *O*-linked glycoproteins, *O*-glycosylated proteins, *O*-glycans, or mucin-type glycoproteins since this is the primary mode of linkage in mucins. The majority of the known glycoproteins contain one or the other type of linkage as the predominant one and therefore do not present a problem in their classification. For example, glycophorin A having 15 oligosaccharides linked to Ser/Thr residues and only one oligosaccharide linked to Asn is clearly an *O*-glycosylated protein. However, there are several glycoproteins in which both linkages are almost equally represented and therefore a clear distinction cannot be easily made. Examples of such *N, O*-glycosylated proteins include fetuin, immunoglobulin A, and human chorionic gonadotropin. However, based on their serum source the former two are generally considered as serum-type, *N*-glycosylated proteins. The nomenclature is further complicated by the fact that other types of *O*-glycosidic linkages between saccharide, albeit nonsialylated, and amino acids exist. These include xylose–serine linkage in proteoglycans, single GlcNAc linked to Ser/Thr in nuclear/cytoplasmic proteins, galactose linked to the δ hydroxyl group of hydroxylysine in collagens, and mannose linked to Ser (Thr) in yeast mannan and arabinose linked to hydroxyproline in plant glycoproteins (Lis and Sharon, 1993). Thus, strictly the term *O*-linked glycoproteins and *O*-glycosylated proteins should include all of the above and the term *O*-glycan should also include polysaccharides such as glycogen, cellulose, and galactans. While an insider is generally able to understand the sometimes confusing terminology, for investigators and students of this field who are not familiar with glycoconjugates this could be problematic and confusing.

## 3. OLIGOSACCHARIDE STRUCTURES OF SIALOGLYCOPROTEINS

The identity of the monosaccharide and amino acid involved in linkage is not the only characteristic by which the classes of sialoglycoproteins could be distinguished. The composition and structure of the saccharides linked to Asn or Ser/Thr are also markedly different. Compositionally the most distinctive feature is that the former always contain Man and seldom contain GalNAc, the latter always contain GalNAc and never Man. Therefore, it is possible to obtain tenta-

tive information on the nature of the sialoglycoprotein by analysis of the monosaccharide composition.

### 3.1. Asn-Linked Saccharides

The saccharide chains linked to Asn show enormous structural variations and are usually classified into three groups (Figure 1) referred to as the high mannose type, complextype, and hybrid type, all of which contain the pentasaccharide Manα1→6(Manα1→3)Manβ1→4GlcNAcβ1→4GlcNAc as a common core structure (boxed in Figure 1a). However, only the complex-type and hybrid-type sugar chains contain siliac acids. High-mannose-type sugar chains contain only Man and GlcNAc residues, and the heterogeneity in these sugar chains is the result of variations in the locations and numbers of Manα1→2 residues linked to the three terminal α-mannose residues of the heptasaccharide: Manα1 → 6(Manα1 → 3)Manα1 → 6(Manα1 → 3)Manβ1 → 4GlcNAcβ1 → 4GlcNAc (Figure 1a).

Complex-type sugar chains exhibit a wide spectrum of structures, since both

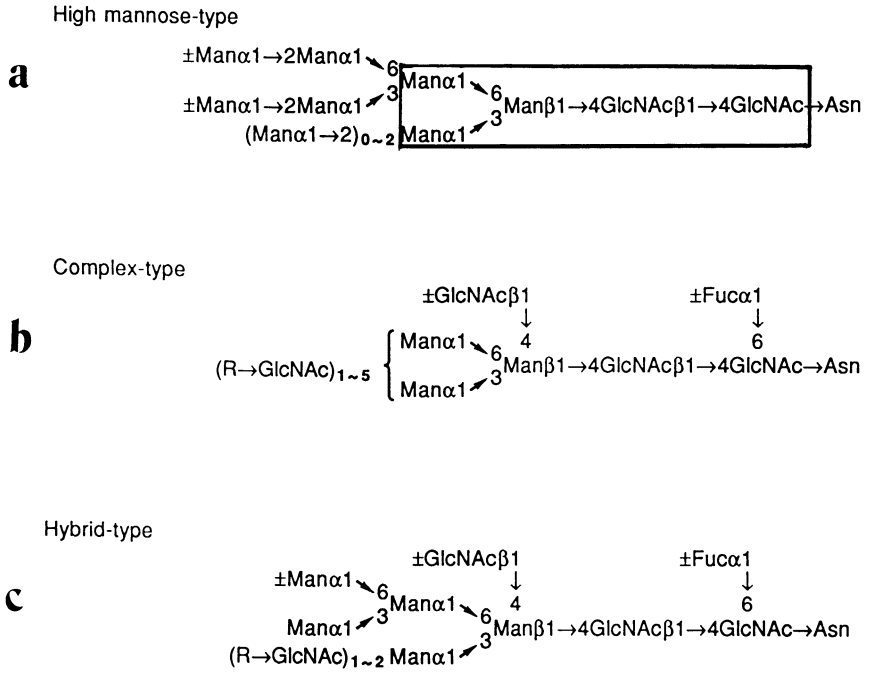
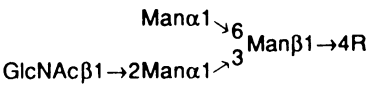


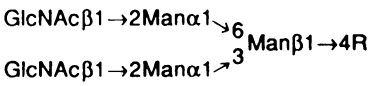

**FIGURE 1.** Three major types of Asn-linked saccharide structures found in glycoproteins. The core pentasaccharide is shown boxed in panel a.

the core and the outer chain moieties exhibit structural variations. Variations in the core include fucosylation of the C-6 position of the proximal GlcNAc residue, and *N*-acetylglucosaminylation of the C-4 position of the β-mannosyl residue; this latter sugar residue is referred to as the bisecting GlcNAc (Figure 1b). The outer chains, which are devoid of mannose, are formed by the addition of up to five residues to the two α-mannosyl residues of the pentasaccharide core, thus resulting in mono-, bi-, tri-, tetra-, and penta-antennary complex-type sugar chains (Figure 2). Next, β-galactose residues are attached to each of the GlcNAc residues to form type 1 chains (Galβ1→3GlcNAc) or type 2 chains (Galβ1→4GlcNAc). The type 2 chain is more widely observed in glycoprotein sugar chains than the type 1 chain. Recently, β-*N*-acetylgalactosamine has been shown to occur in place of the galactose residues as in the *N*-linked sugar chains of most bovine mammary epithelial glycoproteins (Sato *et al.*, 1993b). Finally, variations of the outer chains occur because of differences in the chain termination by the linkage of α-galactose, α-fucose, α-*N*-acetylgalactosamine, sialic acid, and/or sulfate (Figure 3). In the *N*-linked sugar chains, sialic acid residues are most commonly present in α2→3 and α2→6 linkages such as Neu5Acα2→3Gal, Neu5Acα2→6Gal, and Neu5Acα2→6GlcNAc. Glycoproteins obtained from CHO cells and human placenta have sialyl oligosaccharides containing exclusively α2,3-sialyl linkages (Takeuchi *et al.*, 1988; Endo *et al.*, 1988). The Neu5Acα2→6GlcNAc sequence is found in the sugar chains of bovine prothrombin, blood coagulation factor IX, and fetuin expressed solely in the Neu5Acα2→3Galβ1→3GlcNAc branch (Mizuochi *et al.*, 1980). Since the occurrence of the above type 1 chain is limited, th Neu5Acα2→6GlcNAc sequence is rare in glycoproteins. Fucosylation of the Gal and/or the GlcNAc residue of type 1 and type 2 chains by different linkages gives rise to Lewis (Le)-type and ABO-type blood group determinants (Figure 3). Sialylation followed by fucosylation of type 1 and type 2 chains results in formation of the sialyl $Le^a$ and sialyl $Le^x$ structures. In human leukocytes, sialyl $Le^a$ and sialyl $Le^x$ determinants bind selectins expressed on the blood vessel endothelial cells and platelets, causing various cellular interactions and inducing inflammation (Bevilacqua and Nelson, 1993). CD11/CD18 and CD45 are adhesion molecules expressed on leukocytes. Structural studies of the *N*-linked sugar chains of these molecules revealed that CD11/CD18, but not CD45, contain the oligosaccharides with $Le^x$ and sialyl $Le^x$ determinant (Asada *et al.*, 1991; Sato *et al.*, 1993a). The sialyl $Le^x$ determinants of CD11/CD18 were shown to bind to E-selectin *in vitro*, suggesting their involvement in the selectin-mediated adhesion process (Kotovuori *et al.*, 1993). In contrast, CD45 containing exclusively oligosaccharides with the Neu5Acα2→6Gal structures (Sato *et al.*, 1993b), serves as a ligand for CD22, a sialic acid binding lectin, found on the surface of B lymphocytes (Sgroi *et al.*, 1993). It is very interesting that even though the CD11/CD18 and CD45 adhesion molecules are produced by the same cells, the sialylation of these glycoproteins is

1) Monoantennary

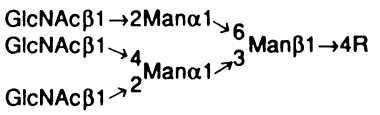



2) Biantennary

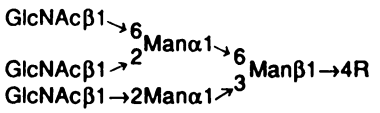



3) Triantennary

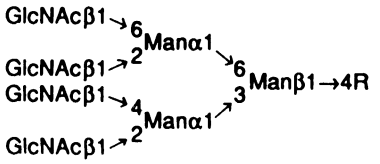
a) 2,4-branched



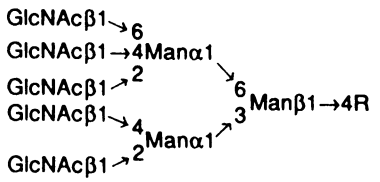

b) 2,6-branched




4) Tetraantennary




5) Pentaantennary




R = GlcNAcβ1→4GlcNAc→Asn

**FIGURE 2.** Branch variations in the complex-type Asn-linked oligosaccharides which give rise to multiantennary structures.

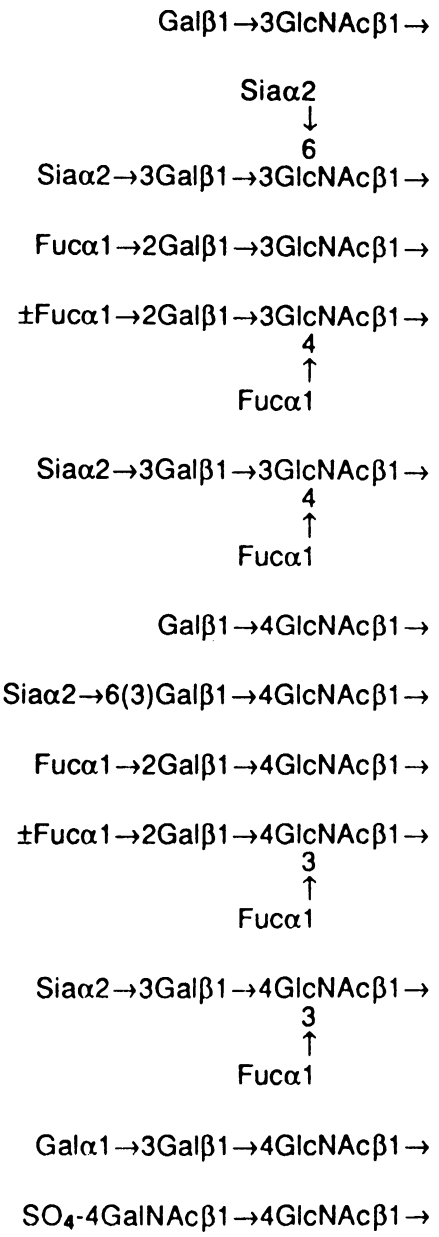



**FIGURE 3.** List of peripheral saccharide structures that terminate the complex-type Asn-linked oligosaccharides of glycoproteins.

very different. Currently, there is no explanation for the differential sialylation of these glycoproteins but it is possible that this is brought about by sialyltransferases whose specificities are influenced by peptide sequences (see Chapter 3). Glycosyltransferases such as the β-*N*-acetylgalactosaminyltransferase in bovine pituitary gland which, in addition to saccharides, also recognize specific peptide sequences in the acceptor molecule are known (see Section 5.2).

The third group of Asn-linked saccharides consist of sugar chains with structural features of both the high mannose type and complex type, and are therefore referred to as the hybrid type (Figure 1c). One or two α-mannose residues are linked to the Manα1→6Man branch of the core, and the outer chain structures of complex type are expressed on the Manα1→3Man side of the core. Recently, a novel hybrid-type sugar chain having the Manα1→Manα1→3Manα1→6Man outer chain has been reported to occur in bovine mammary epithelial glycoproteins (Sato *et al.*, 1993b). The formation of this hybrid-type sugar chain cannot be explained by any of the currently known biosynthetic pathways.

## 3.2. Ser/Thr-Linked Saccharides

A general structure for the saccharides linked to Ser/Thr could be depicted as follows:

| Peripheral | Backbone (not always present) | Core |
|---|---|---|
| Sialic acid, fucose | Gal | Gal, GalNAc |
| Sulfate, GalNAc | GlcNAc | GlcNAc |

The seven different core structures so far found in this class of saccharides are presented in Figure 4. While core structures 1 to 4 show broad species and organ distribution, core structures 5, 6, and 7 have been found so far only in a few glycoproteins such as human rectal adenocarcinoma glycoprotein (Kurosaka *et al.*, 1983), salivary gland mucin of swiftlet (Wieruszeski *et al.*, 1987), and bovine submaxillary mucin (Chai *et al.*, 1992). In a general way, the numbering approximately reflects the frequency of occurrence, core structures 1 and 2 being the most common and core 7 being rare. The simplest sialyl saccharide found in these glycoproteins is the disaccharide SAα2→6GalNAc (the sialyl Tn antigen structure) since sialic acid by itself has not been found directly linked to protein (cf. GlcNAc-Ser/Thr in nuclear/cytoplasmic glycoproteins). It has been proposed that the addition of sialic acid to the linkage GalNAc blocks synthesis of the above core structures and thereby results in the termination of elongation of the saccharide chain. The major portion (~80%) of the carbohydrate in ovine submaxillary mucin exists as this disaccharide (Murthy and Horowitz, 1966). It is clear from the core structures that Ser/Thr-linked saccharides can be unbranched or branched. The *O*-linked saccharide chains are terminated typically,

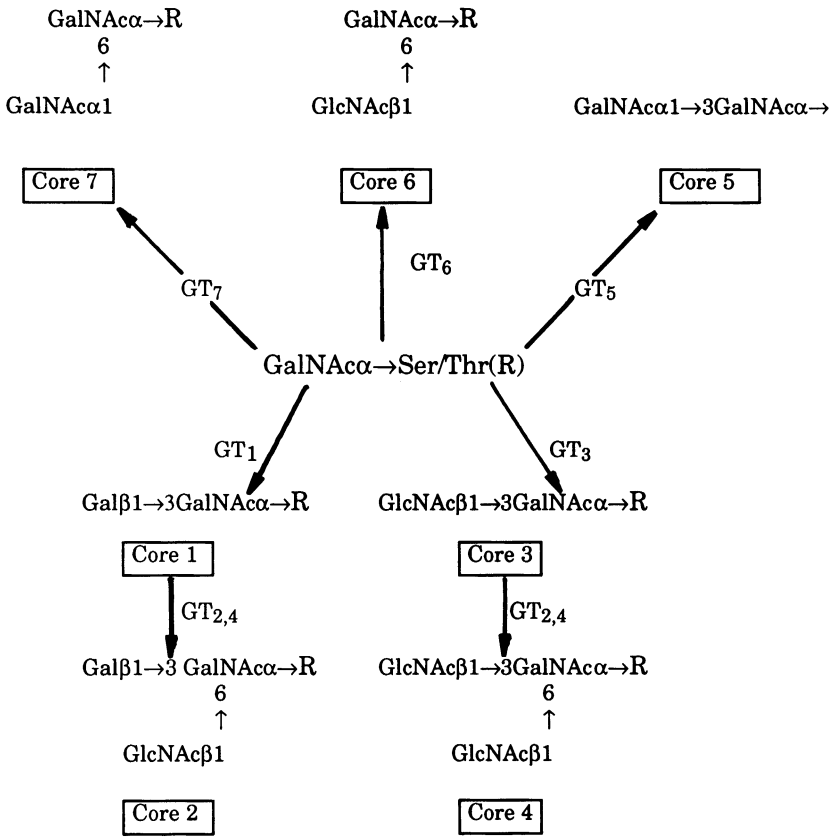



**FIGURE 4.** Biosynthetic pathway for the formation of core structures of Ser/Thr-linked oligosaccharides in glycoproteins. $GT_1$ to $GT_7$ refer to the glycosyltransferases that catalyze the reaction; for example, $GT_1$ is UDP Gal: GalNAcα → R, β1 → 3 galactosyltransferase.

but not exclusively, by addition of sialic acids, fucose, and sulfate residues. Sialic acids and fucose are always linked by α anomeric linkages. Occasionally the saccharide chains are terminated by α-linked Gal or GalNAc. Sulfation when it occurs is on galactose, GlcNAc, or GalNAc residues, either terminal or internal. The combination of different core structures, chain elongation, formation of branch chains, and variation of terminations give rise to a bewildering variety of saccharides in this class of sialoglycoproteins. The peripheral saccharide structures in this class of sialoglycoproteins also include those that confer blood group status such as ABO or Lewis, as discussed above for the Asn-linked saccharides.

The number of oligosaccharides per molecule of protein in this class of glycoproteins range from a few (three?) to a few hundred occurring as dense

clusters along the protein backbone. In fact there is no information on the exact number of oligosaccharide units for any epithelial mucin molecule which usually has a high content (50 to 80% by weight) of carbohydrate compared to other sialoglycoproteins. Further, it appears that unlike most proteins, which are transcribed to yield a homogeneous population, mucin core proteins could consist of variable numbers of repeating sequences (tandem repeats) giving rise to a varying number of glycosylation sites (further discussed below). Thus, the presence of a very heterogeneous population of saccharide units and the variation in the peptide core are probably both equally responsible for the polydisperse nature of mucin preparations noted by both physical and chemical techniques.

There are two clearly distinguishable subclasses of *O*-linked sialoglycoproteins having a high (>50%) carbohydrate content. Group 1 consists of mucins which are the major components of epithelial mucus secretions. These secretory mucins generally have poorly defined, but apparently very high molecular weights (on the order of millions), and in solution yield highly viscous solutions or gels. Group 2 consists primarily of cell surface membrane-associated sialoglycoproteins such as erythrocyte membrane glycophorins (Cartron and Rahuel, 1992), human white blood cell-associated leukosialin (CD43, sialophorin) (Fukuda, 1991a), platelet membrane-associated CD42b (glycocalcin, GP1b) (Okumura *et al.*, 1976), mouse macrosialin (CD68) (Fukuda, 1991a), human milk fat granule membrane glycoproteins (Shimizu and Yamauchi, 1982) and decay-accelerating factor (CD55) of human complement (Medof *et al.*, 1987). In fact, many of the still poorly characterized leukocyte glycoproteins such as the leukocyte common antigen CD45RA, CD96, and PSGL-1 probably belong to this subclass of *O*-linked glycoproteins (Thomas, 1989). Recently, the interest in this subclass of mucin-type glycoprotein has greatly intensified because of the realization that they constitute the primary ligands of selectins. Endothelial cell-associated mucin-type glycoproteins such as GlyCAM-1, CD34, and MAdCAM-1 have all been reported to be ligands for L-selectin (Shimizu and Shaw, 1993, and references therein). The leukocyte membrane mucin-type glycoprotein PSGL-1 has been shown to be a ligand for P-and E-selectin (Sako *et al.*, 1993). The expression of this subclass of membrane-associated cell surface *O*-linked glycoproteins is markedly increased in malignantly transformed cells (see Section 6.2).

Mannose is invariably found as a component of several purified secretory mucins (e.g., see Van Nieuw Amerongen *et al.*, 1987) but because of the difficulties in entirely ruling out the co-purification of small amounts of *N*-glycosylated proteins and in establishing purity of mucin preparations, the significance of this has not been very clear (Strous and Dekker, 1992). In fact, earlier definition of mucins suggested the absence of Asn-linked saccharides and therefore mannose as a distinct feature (Pigman, 1977). The recent detection of potential *N*-glycosylation sites in the deduced amino acid sequence of several mucin core proteins provides stronger evidence for the presence of Asn-linked saccharides as

a true component of mucins (Eckhardt *et al.*, 1991; Xu *et al.*, 1992; Verma and Davidson, 1993). However, to conclude that these sites are actually glycosylated in the final mature product, direct proof is necessary. In contrast to the situation in epithelial mucins (members of group 1 discussed above), the presence of Asn-linked saccharides in membrane-associated mucin-type glycoproteins (members of group 2) such as glycophorin (Tomita and Marchesi, 1975) and leukosialin (Fukuda, 1991a) has been established unequivocally. The biosynthetic studies on MUC1 mucin in breast cells (Hilkens and Buijs, 1988) as well as direct analysis of metabolically labeled MUC1 mucin from H.EP.2 cells (DiIulio and Bhavanandan, unpublished results) have established the presence of Asn-linked saccharide in this membrane-associated mucin-type glycoprotein as well.

### 3.3. Polylactosaminoglycans

The glycosaminoglycan keratan sulfate consists of long chains of repeating disaccharide units made up of galactose linked via β→4 to GlcNAc either of which may be sulfated in position 6 (Bhavanandan and Meyer, 1966). Two types of keratan sulfate chains can be distinguished based primarily on the nature of the linkage of the polysaccharide chains to the protein. In corneal keratan sulfate, the polysaccharide chains are linked to Asn, and in skeletal keratan sulfate, they are linked to Ser/Thr. Similar polysaccharide chains of repeating *N*-acetyllactosamine units commonly referred to as polylactosaminoglycans are also found in various glycoproteins as part of either *N*-or *O*-linked sugar chains and in glycolipids. The polylactosaminoglycan chain can be either linear or branched. In the linear form, the *N*-acetyllactosamine structures are linked to each other by β-1,3 linkages, while in the branched form the C-6 position of some galactose residues in the repeating units are substituted with additional *N*-acetyllactosamine. These linear and branched structures constitute the blood group i and I antigenic determinants, respectively (Watanabe *et al.*, 1979).

In the Asn-linked glycoproteins, the polylactosaminoglycan chains are attached to the trimannosyl core and show variations in the number and location of the repeating units. The backbone of polylactosaminoglycans consists of the type 2 Galβ1→GlcNAc structure, but at the distal portions of the chains Galβ1→3-GlcNAc structures of type 1 are often found. ABH and Lewis blood group antigenic determinants are usually expressed on the type 1 and type 2 chains attached directly to the trimannosyl core. The nonreducing terminal galactose residues are usually sialylated and most of the sialyl polylactosaminoglycans contain the Neu5Acα2→3Gal group, even though Neu5Acα2→Gal groups are also known to occur.

In a variety of cells, polylactosaminoglycans are mainly carried by lysosomal membrane glycoproteins, LAMP-1 and LAMP-2 (Do *et al.*, 1990; Carlsson and Fukuda, 1992), portions of which are located on cell surfaces (Lippincott-

Schwartz and Fambrough, 1987; Carlsson and Fukuda, 1992). Since the terminal structures of polylactosaminoglycans can express the sialyl $Le^x$ and sialyl $Le^a$ determinants, which are believed to be the natural ligands involved in selectin-mediated cell adhesion, the expression levels of LAMPs may be biologically important.

## 4. PROTEIN CORE OF *O*-LINKED SIALOGLYCOPROTEINS

The information available on the nature of the core proteins of the *O*-glycosylated proteins is still sketchy and the true sizes of the protein backbone of epithelial mucins are unknown (Strous and Dekker, 1992). Prior to the 1980s there was hardly any information concerning the core proteins of this class of glycoconjugates because of difficulties in deglycosylating and in applying the classical techniques of peptide sequence determination. The problem was particularly serious in the case of the epithelial mucins because of the apparent large size of the native molecules and estimated protein core size in excess of 200,000. Thus, usually only partial peptide sequences could be obtained; for example, in the case of ovine submaxillary mucin, the sequence of only 62 amino acids out of an estimated total 650 amino acids was obtained by analyzing tryptic peptides (Hill *et al.*, 1977a,b). The situation was somewhat better for the lower-molecular-weight, membrane-associated, members of this class of glycoconjugates. The entire peptide sequence of glycophorin was reported by Tomita and Marchesi (1975). This sequence, for the first time, established the presence of a stretch of hydrophobic amino acids of sufficient length to traverse the lipid bilayer. The sequence also confirmed a feature that has been predicted for *O*-glycosylated protein, specifically mucins, but not proven, namely, the presence of consecutively glycosylated serine and threonine residues. However, two recent studies found that the previously proposed cluster of six glycosylation sites in glycophorin is an error and that only four of these are occupied (Nakada *et al.*, 1993; Pisano *et al.*, 1993). However, the clustered saccharide domains provide a dense array of clustered *O*-linked saccharides presumably with important functional implications. Structurally the presence of a high density of sialylated and sometimes sulfated, and therefore negatively charged saccharides results in an extended conformation of the molecule (Gottschalk, 1960; Gottschalk and Thomas, 1961). Interestingly, it appears that the removal of sialic acid and hence the charge does not dramatically alter the conformation, as was previously believed, and the molecules with only a high density of single GalNAc residues still retain the extended rod shape (Shogren *et al.*, 1989). On the cell surface these extended *O*-linked glycoproteins would provide multiple saccharide binding sites for specific receptors. The ligands for lectins, antibodies, and selectins in many instances are likely to be multivalent saccharide structures. This was previously

demonstrated to be the case for the lectin wheat germ agglutinin (Bhavanandan and Katlic, 1979; Furukawa *et al.*, 1986). This lectin strongly interacts with glycophorin T1 glycopeptide containing three sialylated saccharides on consecutive amino acids, but does not bind to this same saccharide when present individually as in glycopeptides isolated from fetuin, or if released from the T1 glycopeptide by β-elimination. Similarly, it has been demonstrated that three GalNAc residues present as a cluster on the peptide sequence SerThrThr are essential for recognition and binding by the MLS 128 monoclonal antibody (Nakada *et al.*, 1991).

Sequences of several other membrane-associated heavily *O*-glycosylated proteins have since been obtained. A notable one is that of leukosialin (CD43) having 385 amino acids, of which 239 constitute the mucinlike extracellular domain estimated to carry between 70 and 85 Ser/Thr-linked oligosaccharides (Fukuda, 1991a). A new feature not present in glycophorin is the tandem repeat of amino acid sequences. In leukosialin four repeats of 18 amino acid sequences are present. Tandemly repeated sequences are now considered a standard feature of mucins and mucin-type glycoproteins; however, it may be premature to make a definite conclusion in view of the limited number of full-length core peptide sequences available.

Amino acid analysis of mucin proteins reveals high levels of serine, threonine, proline, glycine, alanine, aspartic acid, glutamic acid, and low levels of aromatic and sulfur-containing amino acids. Chemical and enzymatic deglycosylation have yielded low values for molecular masses, ranging from about 60 kDa to >200 kDa for submaxillary, tracheal, and breast/pancreatic tumor MUC1 mucins (Hill *et al.*, 1977a; Burchell *et al.*, 1987; Bhavanandan and Hegarty, 1987; Woodward *et al.*, 1987; Lan *et al.*, 1990). In contrast, some of the estimates based on *in vitro* translation and biosynthetic studies indicate higher values of 160 kDa for human intestinal mucin (Gum *et al.*, 1989) and MUC1 mucin (Hilkens and Buijs, 1988), 400 kDa for human bronchial mucin (Marianne *et al.*, 1987), and 900 kDa for human gastric mucin (Dekker *et al.*, 1991). The discrepancies are not fully explicable at present but the lower values obtained by deglycosylation are believed to be related to breakage of peptide bonds followed probably by amino-terminal modification since usually no new amino-termini are detectable after deglycosylation (Hill *et al.*, 1977a; Bhavanandan and Hegarty, 1987; Woodward *et al.*, 1987).

Recent investigations using molecular biological techniques have provided new insights concerning the core protein of mucins. However, the information is still fragmentary since the majority of the deduced amino acid sequences are from partial cDNA clones. Currently, only a few full-length mucin glycoprotein sequences deduced from cDNA clones have been reported. These include the human and mouse MUC1 mucin (Gendler *et al.*, 1990; Ligtenberg *et al.*, 1990; Spicer *et al.*, 1991), the frog integumentary mucin glycoprotein FIM (Probst *et*

*al.*, 1992), a low-molecular-weight human salivary mucin (Bobek *et al.*, 1993), and canine tracheobronchial mucin (Verma and Davidson, 1993). The cDNA sequence of a bovine submaxillary mucinlike protein previously reported to be full length (Bhargava *et al.*, 1990) was subsequently found to be incomplete (Woitach, Kiel, and Bhavanandan, unpublished results). A number of other partial cDNAs of mucin core proteins have been sequenced revealing several characteristic features. A feature noted in many, but not all, core proteins is a variable number of tandem repeats (Table I). These repeats, located in the central region of the molecule, vary in size from 6 to 169 and are enriched in serine and threonine and small amino acid residues such as proline, glycine, and alanine. It appears that these repeats provide the attachment sites for the bulk of the Ser/Thr-linked saccharides in the molecule, but in the human salivary MG2 mucin almost half of the carbohydrate is estimated to be located outside of the tandem repeat region (Bobek *et al.*, 1993). The deduced (partial or full) peptide sequences which have so far not revealed a tandem repeat feature include bovine submaxillary mucin (Bhargava *et al.*, 1990; Woitach, J. W., Keil, R., and Bhavanandan, V. P., unpublished results), canine tracheobronchial mucin (Verma and Davidson, 1993), rat intestinal mucin (Xu *et al.*, 1992), and certain clones of human tracheobronchial mucin (Dufosse *et al.*, 1993). The regions outside of the tandem repeats, constituting the N- and C-terminal ends of the core protein, have a wider representation of amino acids. Many of the C-terminal (Bhargava *et al.*, 1990; Eckhardt *et al.*, 1991; Xu *et al.*, 1992) and a few N-terminal segments (Gum *et al.*, 1993; Probst *et al.*, 1992) are enriched in cysteine, an unexpected feature since, typically, purified mucins have very little or no cysteine. In one case, that of human small intestinal MUC2 mucin, a cysteine-rich segment is located in the interior of the molecule, where it separates tandem repeats (Toribara *et al.*, 1991; Gum *et al.*, 1993). The regions outside of the tandem repeats also contain several potential sites for *N*-glycosylation; there are none in the tandem repeat segments. Biosynthetic studies indicate that *N*-glycosylation actually occurs, and precedes *O*-glycosylation of rat gastric mucin (Dekker and Strous, 1990). Since the subsequent processing events have not been fully elucidated, whether these saccharides are removed by processing or remain in the final mature mucin molecule is still not clear. Finally, putative transmembrane domains are present in peptide sequences of membrane-associated *O*-glycosylated proteins such as leukosialin and MUC1 glycoprotein, but not in the sequences of secreted mucins which have been deduced to date.

At present, seven different human mucin glycoprotein genes, termed MUC1 to MUC7, have been described (Table I). Of these the full-length sequence for the MUC1 mucin associated with breast, pancreatic, laryngeal, and other carcinomas is of special interest. The sequence reveals 40 to 90 tandem repeats of 20 amino acids each and additional repeats which differ in a few amino acids. Two signal sequence variants were detected in the N-terminus segment indicating

**Table I**
**Tandemly Repeated Amino Acid Sequences Found in *O*-Glycosylated Proteins**

| | | References |
|---|---|---|
| Human *O*-glycosylated proteins | | |
| Glycophorin[a] | None | Tomita and Marchesi (1975) |
| Leukosialin[a] | TSGPPVTMATDSLETSTG | Fukuda (1991a) |
| MUC1[a] | PDTRPAPGSTAPPAHGVTSA | Gendler *et al.* (1990), Ligtenberg *et al.* (1990), Williams *et al.* (1990) |
| MUC2[a] | PTTTPITTTTTVTPTPTPTGTQT | Toribara *et al.* (1991) |
| MUC3 | HSTPSFTSSITTTETTS | Gum *et al.* (1990) |
| MUC4 | TSSASTGHATPLPVTD | Porchet *et al.* (1991) |
| MUC5 | TTSTTSAP | Crepin *et al.* (1990) |
| MUC6 | SPFSSTGPMTATSFQTTTTYPTPSHPOTTLPT HVPPFSTSLVTPSTGTVITPTHAQMATSASH STPTGTIPPPTTLKATGSTHTAPPMTPTTSTS QAHSSFSTAKTSTSLHSHTSSTHHPEVTSTTT ITPNPTSTGTSTPVAHTTSATSSRLPTPFTTH SPPTGS | Toribara *et al.* (1993) |
| MUC7 | TTAAPPTPSATTPAPPSSSAPPE | Bobek *et al.* (1993) |
| Nonhuman *O*-glycosylated proteins | | |
| Mouse MUC1[a] | PATSAPXDSTSSPVHSGTSS | Spicer *et al.* (1991) |
| Rat intestinal M-2 (RMUC 176) | TTTPDV | Khatri *et al.* (1993) |
| Rat intestinal MLP (mucinlike protein) | PSTPSTPPPST | Xu *et al.* (1992) |
| Rat airway mucinlike protein | TTTTIITI | Tsuda *et al.* (1993) |
| Frog integumentary mucin-A·1 (spasmolysin) | VPTTPETTT | Hoffmann (1988) |
| Frog integumentary mucin-B·1 | GESTPAPSETT | Probst *et al.* (1992) |
| Porcine submaxillary mucin | GAGPGTTASSVGVTETARPS VAGSTTGTVSGASGSTGSSSG SPGATGASIGPETSRISVAGSS GAPAVSSGASQATS | Timpte *et al.* (1988) |
| Bovine submaxillary mucin | None | Bhargava *et al.* (1990) |
| Canine tracheobronchial mucin[a] | None | Verma and Davidson (1993) |

[a]Full-length sequences available; all others are partial sequences deduced from cDNA.

alternate splicing. The C-terminus segment contains a 28-amino-acid hydrophobic region and 68 amino acids beyond the putative transmembrane segment which probably constitutes the cytoplasmic tail. Five potential *N*-glycosylation sites are also present in the nonrepetitive C-terminal segment. Five potential *O*-glycosylation sites are present within each tandem repeat of 20 amino acids and an equal proportion of additional sites in the nonrepetitive domains. However, there is no information as to which of these sites are occupied in the mature MUC1 mucin. The other six listed in Table I are tandem repeats from partial cDNA clones coding for secreted epithelial mucins. The human intestinal mucin coded by the MUC2 gene consists of 51 to 115 tandem repeats of 23 amino acids in the middle, with additional incompletely conserved repeats at the *N*-terminal end (Toribara *et al.*, 1991). Multiple cysteine residues, including several Cys-Cys dipeptides, are present in the sequence but not all of these are confined to the C-terminal segment. The N-terminal segment has a number of potential *N*-glycosylation sites and most of the aromatic amino acids. A different human intestinal mucin core protein coded by the MUC3 gene has tandem repeats of 17 amino acids. It is proposed that this core protein predominates in the acidic intestinal mucin while the MUC2 core protein predominates in the neutral intestinal mucin (Gum *et al.*, 1990). The MUC4 and MUC5 genes code for human tracheobronchial mucin core proteins with 16 and 8 amino acid tandem repeats, respectively. In the case of MUC4, 38 repeats are present but only 8 of the 16 amino acids are perfectly conserved in them (Porchet *et al.*, 1991). MUC6 codes for a gastric mucin core protein which has the longest repeat, of 169 amino acids, so far reported (Toribara *et al.*, 1993). MUC7 codes for a small salivary mucin glycoprotein of 120–150 kDa and the central region of the core protein consists of six perfect tandem repeats of 23 amino acids of which 9 are Thr and Ser (Bobek et al, 1993). A putative signal peptide of 20 hydrophobic amino acids in the N-terminal end and five potential *N*-glycosylation sites are also present in the MUC7 sequence. The tandem repeats of these seven human mucin proteins, varying between 8 and 169 amino acid residues, all have high proportions of hydroxy amino acids but there is no sequence homology at either the amino acid or nucleotide levels. In fact the only other common amino acids in all of them are proline and alanine. A point of interest is that while all reported tandem repeats have potential *O*-glycosylation sites, there is a preponderance of threonine in some with the MUC2 tandem repeat having only threonine. The significance of this is unclear at the present, but the possibility that mucins may be coded by more than one gene should be considered. In fact there is already evidence for the presence of more than one cDNA clone in human tracheobronchial (Dufosse *et al.*, 1993), rat intestinal (Khatri *et al.*, 1993), and bovine submaxillary mucin (Woitach, Keil, and Bhavanandan, unpublished results) cDNA libraries. The lack of sequence homology at either the nucleotide or amino acid levels between even closely related molecules such as human and rat gastric mucins or human

and mouse MUC1 mucin-type glycoproteins is surprising. So far, the greatest sequence homology in mucin core proteins has been noted in the cysteine-rich C-terminal domains outside of the tandem repeats. There is an 82% sequence similarity at the protein level including exact matches of 30 cysteine residues in the C-terminal domains of porcine and bovine submaxillary mucins (Bhargava *et al.*, 1990; Eckhardt *et al.*, 1991). The amino acid homology in the transmembrane and C-terminal domains of the human and mouse MUC1 mucin is 87% (Spicer *et al.*, 1991). There are also lesser amino acid sequence similarities between the C-terminal domains of the bovine and porcine submaxillary mucins, MUC2 mucin, and frog integumentary mucin B.1 (Probst *et al.*, 1992). One other sequence similarity is the 73% homology between the C-terminal 700 amino acids of human intestinal MUC2 mucin and the rat intestinal mucinlike peptide (MLP) (Xu *et al.*, 1992). The high degree of sequence conservation in the C-terminus of both the secretory and membrane-associated mucins suggests some functional importance. However, as mentioned above, most purified secretory mucins are poor in cysteine, and therefore it is necessary to clearly demonstrate first that the cysteine-rich domains are intact in the fully processed mature mucin. The role of disulfide interactions in the polymeric structures of secretory mucins has been controversial. The rheological properties of most, but not all, mucins appear to be dependent on S-S bridge formation. Eckhardt *et al.* (1991) have suggested that in the case of the porcine submaxillary mucin the cysteine-rich segments are lost during purification as a result of proteolysis. However, the cysteine content of bovine submaxillary mucin prepared under stringent conditions, to avoid proteolysis, does not differ significantly from that of normal preparations (Woitach and Bhavanandan, unpublished results). For the MUC1 mucins the conserved cytoplasmic domain is thought to be important in linking the transmembrane glycoprotein to cytoskeletal elements such as actin (Spicer *et al.*, 1991). A striking difference between the secretory mucins and the transmembrane *O*-glycosylated proteins (glycophorin, leukosialin, MUC1 mucin, CD 34) is the lack of cysteine residues in the C-terminal regions (cytoplasmic tail) of the latter. This suggests that intra- and/or intermolecular disulfide bonds are not essential for the function of the membrane-associated mucin-type glycoproteins.

Not all mucin core protein sequences known at present have tandem repeats. The mucin sequences that lack tandem repeats, however, do have repetitive peptide motifs. Thus, the partial sequence of bovine submaxillary mucin has repeats of the sequence GTTVAPGSSNT (Bhargava *et al.*, 1990) and canine tracheobronchial mucin contains the motifs TPTPTP and TTTTPV repeated 13 and 19 times, respectively (Verma and Davidson, 1993). Dufosse *et al.* (1993) recently reported a family of cDNA clones isolated from a human tracheobronchial library showing no tandem repeats but a new type of peptide organization. These clones contain degenerate 87-base-pair tandem repeats which encode for nonrepetitive peptide sequences. A new feature noted in these deduced se-

quences, which lack tandem repeats, is the distinct alternating hydrophilic and hydrophobic domains.

The loci of several mucin genes have now been mapped. MUC1 has been mapped in the region of 22q on chromosome 1 (Swallow *et al.*, 1987a), MUC2 to chromosome 11 p15 (Griffiths *et al.*, 1990), MUC3 to chromosome 7 (Gum *et al.*, 1990), and MUC4 to chromosome 3 (Porchet *et al.*, 1991). Thus, the mucin genes are not clustered in the genome.

## 5. BIOSYNTHESIS OF SIALOGLYCOPROTEINS

### 5.1. General Concepts

The protein backbones of all sialylated glycoproteins are synthesized on polyribosomes on the cytoplasmic face of rough endoplasmic reticulum and cotranslationally translocated into the lumen where glycosylation can occur. The incorporation of glycosyl residues into the polypeptide chain requires the presence of the appropriate glycosyltransferases and the donor sugar–nucleotide. The general reaction catalyzed by glycosyltransferase is

$$\text{Sugar–Nucleotide} + \text{Acceptor} \rightarrow \text{Sugar–Acceptor} + \text{Nucleotide}$$

The glycosyltransferases involved in glycoconjugate biosynthesis are membrane-bound enzymes and when purified require detergents for optimal activity. The glycosyltransferase has an absolute donor (sugar–nucleotide) specificity and the name, e.g., galactosyltransferase, indicates this specificity, i.e, the requirement for UDP-Gal. The glycosyltransferase also has specificity for the acceptor to which it will transfer the sugar and for the type of linkages (e.g., $\alpha 1 \rightarrow 3$, $\beta 1 \rightarrow 4$) that will be formed. Therefore, it is the specificity of the glycosyltransferase that primarily directs and controls the formation of a particular structure. In general, one glycosyltransferase is required for every type of sugar linkage (e.g., sugar $\beta 1 \rightarrow$ x R, where x is position on the acceptor R) that is present in a glycoconjugate. There is no direct genetic control of the synthesis of saccharide structures but indirect control is exerted via expression of specific glycosyltransferase activities as is well illustrated in the synthesis of the blood group ABO(H) determinants. Occasional exceptions to the above rule are known; thus, an external factor may influence/regulate synthesis, such as the requirement for lactalbumin in the formation of the lactose (Gal$\beta 1 \rightarrow$4Glc linkage). In some instances one enzyme may catalyze the formation of two different linkages, while in others there appears to be more than one enzyme for the synthesis of the same linkage, e.g., SA$\alpha 2 \rightarrow$6GalNAc (Kurosawa *et al.*, 1994). Several glycosyltransferases have been cloned and studied at the genetic level. The amino acid sequences reveal common structural domains in these enzymes but only limited

amino acid sequence homology. The glycosyltransferases specifying blood group A or B status differ in a few single-base substitutions which cause changes in four amino acid residues (Yamamoto *et al.*, 1990). The activated sugar nucleotides involved are derivatives of either uridine or guanidine diphosphates, with the exception of sialic acid which is activated as the cytosine monophosphate derivative. All of the sugar nucleotides are biosynthesized on the cytoplasmic side of the membrane and therefore, to be useful, must be transported to the lumen of RER and Golgi, where the glycosyltransferases are located. Recent research has shown that this translocation is facilitated by transporter proteins (antiports) present in the membranes which exchange free nucleotide with sugar–nucleotide on a one-to-one basis (e.g., one molecule of UMP for one UDP-Gal). The details of the regulation of glycoprotein biosynthesis remain to be elucidated. Some specific observations that need explanation include the following:

1. The basis for the existence of extreme diversity of both Asn- and Ser/Thr-saccharides. In the case of the Asn-linked saccharides since all of the structures are derived from the common precursor, $Glc_3Man_9GlcNAc_2$-Asn, there must be fine-tuning at the processing level. The peptide sequence and overall protein structure/conformation are believed to play major roles in this.
2. The factors that determine the glycosylation of the same protein in different animal species or in different organ/tissues of the same animal.
3. The differences noted in the glycosylation patterns in different proteins in the same cell.
4. The variations (heterogeneity) of the saccharides at a single glycosylation site in a protein as found in ovalbumin (the so-called glyco forms).

## 5.2. Biosynthesis of Glycoproteins Containing Asn-Linked Sugar Chains

Since the discovery of the involvement of lipid-linked oligosaccharide, the biosynthesis of Asn-linked sugar chains has been extensively studied (Kornfeld, 1982; Beyer and Hill, 1982). In the biosynthesis of Asn-linked sugar chains, a precursor oligosaccharide consisting of $Glc_3Man_9GlcNAc_2$ is first assembled on a lipid, dolichol, and then transferred to polypeptides by oligosaccharidyltransferase in the rough endoplasmic reticulum. The asparagine residue in the Asn-Xaa-Ser/Thr, in which Xaa can be any amino acid except proline, are potential sites for glycosylation. Since not all of the sites with this sequence are glycosylated, this tripeptide sequence is not sufficient for protein glycosylation. It is suggested that the accessibility of the sequence, in folded peptides, to the oligosaccharidyltransferase may be important. The three glucose residues are cleaved off from the peptide-bound oligosaccharide by the sequential action of α-glucosidases I and II, respectively, resulting in the high-mannose structure, $Man_9GlcNAc_2$. After translocation of glycoproteins to the Golgi apparatus, the

oligosaccharide is further processed to the $Man_5GlcNAc_2$ structure by the action of an α-1,2-mannosidase (α-mannosidase I) which cleaves four α-1,2-linked mannose residues. At this stage a GlcNAc residue may be transferred to the Manα1→3Man arm of the heptasaccharide by *N*-acetylglucosaminyltransferase I (GlcNAcTI), leading to hybrid-type sugar chains. Alternatively, two mannose residues may be removed by the action of a second Golgi apparatus-associated mannosidase (α-mannosidase II), and a GlcNAc residue transferred to the Manα1→6Man arm by GlcNAcTII, leading to the complex-type sugar chains. To these biantennary sugar chains another GlcNAc residue may be added via β-1,4 or β-1,6 linkages by the action of GlcNAcTVI or GlcNAcTV (Cummings *et al.*, 1982), forming 2,4- or 2,6-branched triantennary sugar chains. Tetraantennary sugar chains are produced by the further addition of GlcNAc to the triantennary sugar chains by GlcNAcTVI or GlcNAcTV. Action of GlcNAcTIII and α-1,6-fucosyltransferase on hybrid-type and complex-type sugar chains results in the formation of bisected sugar chains and the fucosylated core, respectively.

In the next step, galactose is transferred to each of the GlcNAc residues, except for bisecting GlcNAc, of the complex-type sugar chains by β-1,3-galactosyltransferase or β-1,4-galactosyltransferase to form type 1 or type 2 structures discussed above. Although the type 1 structure has been found in several glycoproteins, the β-1,3-galactosyltransferase activity responsible for the synthesis of this structure has not been detected. Recently a new variation was found in the Asn-linked sugar chains in which GalNAc instead of Gal is added to the GlcNAc residues. The sequence $SO_4$→4GalNAcβ1→4GlcNAc in Asn-liked saccharides was first detected in mammalian glycohormones (Green *et al.*, 1985) and later in many other glycoproteins (Sato *et al.*, 1993b, and references therein). This unusual saccharide structure is apparently important for the clearance of these glycoproteins from the circulation since it is the ligand for a lectin present on hepatic reticuloendothelial cells (Fiete *et al.*, 1991). In bovine pituitary gland, *N*-acetylgalactosaminylation of glycohormone sugar chains is catalyzed by a β-*N*-acetylgalactosaminyltransferase, which in addition to the GlcNAcβ→-Man→ also recognizes the Pro-Xaa-Arg/Lys sequences located six to nine amino acids away from the N-terminus of putative glycosylation sites (Smith and Baenziger, 1992). However, human urinary kallidinogenase, which is synthesized in the kidney and contains sugar chains with the GalNAcβ1→4GlcNAc structures (Tomiya *et al.*, 1993), lacks the proposed tripeptide recognition motif. This suggests the presence of other β-1,4-*N*-acetylgalactosaminyltransferases with different acceptor specificities.

Finally, the galactose residues are sialylated by sialyltransferases, resulting in the formation of Neu5Acα2→6Gal and Neu5Acα2→3Gal structures. A detailed discussion of sialyltransferases can be found in Chapter 3. An α-2,6-sialyltransferase purified from rat liver can transfer sialic acid to the Galβ1→4GlcNAc group but not to the Galβ1→3GlcNAc group, while α-2,3-sialyltransferase from the same source can transfer sialic acid to type 1 and type 2 chains (Weinstein *et*

*al.*, 1982). Kinetic studies revealed that the α-2,6-sialyltransferase transfers sialic acid to the galactose residue of the Galβ1→4GlcNAcβ1→2Manα1→3Man branches, while α-2,3-sialyltransferase transfers sialic acid to the galactose residue of the Galβ1→4GlcNAcβ1→2Manα1→6Man branches (Joziasse *et al.*, 1987). The cDNAs that encode these sialyltransferases have been cloned (Weinstein *et al.*, 1987; Wen *et al.*, 1992). Since the *in vitro* assay of the α-2,3-sialyltransferase activity showed that the Galβ1→3GlcNAcβ→ sequence is the preferred acceptor compared to the Galβ1→4GlcNAcβ→ sequence, it is assumed that there is another α-2,3-sialyltransferase which more efficiently catalyzes the synthesis of the Neu5Acα2→3Galβ1→4GlcNAc structure (Weinstein *et al.*, 1982). In accordance with this assumption, the cDNA that encodes the α-2,3-sialyltransferase with an acceptor specificity different from that described above has been cloned from a human melanoma cell cDNA library and shown to be involved in the synthesis of the sialyl $Le^x$ determinant in Namalwa cells by gene transfection experiments (Sasaki *et al.*, 1993).

Thus, specific carbohydrate structures are synthesized by the combined action of glycosyltransferases, each of which has a strict acceptor specificity and therefore determines the structure of the product. Differences in expression of the glycosyltransferases in cells result in differences in the carbohydrate structures produced by different cells. Systematic structural analysis of the Asn-linked sugar chains of γ-glutamyltranspeptidase (γ-GTPs) purified from rat, bovine, mouse, and human liver and kidney revealed that the major sugar chains of all kidney enzymes contain bisecting GlcNAc (for review, see Furukawa and Kobata, 1992). The liver enzymes from the four species are devoid of these structures. Further, the sugar chains of mouse and human kidney but not mouse and human liver γ-GTPs contained $Le^x$ antigenic determinant. These findings strongly suggest that the tissue- and species-specific glycosylation is mostly related to the differential expression of the transferases. In other studies, the Asn-linked sugar chains of the blood coagulation factor XIII purified from human plasma and recombinant BHK cells were examined in detail (Hironaka *et al.*, 1992). Both factor VIII preparations contained high-mannose-type and bi-, tri-, and tetraantennary complex-type sugar chains in similar ratios but the outer chain structures were different. Some of the complex-type sugar chains in human plasma factor VIII contained blood group A and/or H determinant, while in the recombinant product these were absent. In addition, a small proportion, less than 10%, of the complex-type sugar chains of recombinant factor VIII had the Galα1→3Gal structure which was not detected in the sugar chains of the human plasma factor VIII. The expression of the Galα1→3Gal group is known to be species-specific and occurs in New World monkeys and nonprimates but not in humans or Old World monkeys (Galili *et al.*, 1988). The cDNA that encodes α-1,3-galactosyltransferase has been cloned from a bovine cDNA library (Joziasse *et al.*, 1989) and from a mouse F9 cell cDNA library (Larsen *et al.*, 1989).

Two genomic DNAs, homologous to the bovine cDNA, have been isolated from a human genomic library but these turned out to be nonfunctional pseudogenes (Larsen *et al.*, 1990; Joziasse *et al.*, 1991), confirming the absence of the Gal-α1→3Gal determinants in human glycoproteins. Thus, the expression of carbohydrate structures is directly controlled by the gene expression of individual glycosyltransferases.

### 5.3. Biosynthesis of Glycoproteins Containing Ser/Thr-Linked Sugar Chains

In contrast to the biosynthesis of Asn-linked saccharides, all of the steps in the synthesis of Ser/Thr-linked saccharides appear to occur by transfer of individual monosaccharides directly from their sugar nucleotide. Thus, all *O*-linked chains in mucin glycoproteins are initiated by the transfer of single GalNAcα from UDP-GalNAc to serine or threonine residues in the polypeptide chain. The reaction is catalyzed by polypeptide GalNAc transferase (UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferase]. Two important issues concerning this initial step still remain to be fully resolved: (1) what factor(s) determine which serine and threonine residues in a protein are glycosylated?, and (2) what is the subcellular location of the polypeptide:GalNAc transferase and when does chain initiation occur? There is no consensus on either issue. Concerning the first issue, extensive research utilizing apomucins and their proteolytic fragments as well as synthetic peptide has failed to reveal a consensus primary amino acid sequence surrounding the glycosylated hydroxy amino acid (Aubert *et al.*, 1976; Hill *et al.*, 1977b; Briand *et al.*, 1981). The available data indicate that peptides containing clusters of serine/threonine residues and those containing proline residues in the vicinity of serine/threonine such as at "−1" and "+3" positions are preferred substrates (Wilson *et al.*, 1991). However, the absence of proline close to 14 out of the 15 glycosylated serine or threonine residues in glycophorin rules out the need for proline in the neighborhood of Ser/Thr as an absolute requirement. Nonetheless, the presence of proline in the vicinity may confer a favorable peptide conformation and presentation of serine/threonine for *O*-glycosylation since it is known that proline residues confer a rigid rodlike conformation on peptides. A serine or threonine residue next to an already *N*-acetylgalactosaminylated serine/threonine may be similarly preferred as a substrate for new chain initiation as a result of the conformational change that had occurred by the addition of the first GalNAc. The attachment of GalNAc residues to peptides is known to alter certain polypeptides from a globular to an extended, rigid conformation (Shogren *et al.*, 1989; Gerken *et al.*, 1989). Based on the corrected amino acid sequence of glycophorin, an attempt was made to formulate rules to determine which serine/threonine in protein would be glycosylated (Pisano *et al.*, 1993). However, the generality of the findings is doubtful since unlike gly-

cophorin, all but one of the tandem repeat amino acid sequences found in mucin glycoproteins so far are rich in proline, the exception being the TTTTIITI repeat in rat airway mucin (Table I). Recent studies on porcine submaxillary gland UDP-GalNAc:polypeptide *N*-acetylgalactosaminyl transferase suggest that the specificity of the enzyme is influenced by the amino acid sequences adjacent to serine or threonine residues to be glycosylated (Wang *et al.*, 1993). In summary, the indications are that even though there is no consensus primary amino acid sequence for *O*-glycosylation comparable to that for *N*-glycosylation, there may be preferences based on signals governed by peptide conformation.

Concerning the timing and localization of the initiation of *O*-glycosylation, again there is no consensus. However, the majority of available evidence indicates that the bulk of the initiation occurs late; that is, after the release of the nascent peptide from the polysomes. Strong evidence in support of this is the localization of the polypeptide GalNAc transferase in the Golgi apparatus, with no detectable activity in rough endoplasmic reticulum, and the demonstration of a transport protein for UDP-GalNAc in Golgi vesicles (Abeijon and Hirschberg, 1987; Roth, 1984; Schweizer *et al.*, 1994). Additionally, it has been noted that less than 1.3% of the potential *O*-linked chains have been initiated in isolated fetuin peptidyl-RNA (Johnson and Heath, 1986). Initiation of *O*-glycosylation of leukosialin was shown to occur only in an early Golgi compartment; the initiated chains were then very rapidly elongated (Piller *et al.*, 1989). Studies on *O*-glycosylated proteins of viruses also confirm that the addition of linkage GalNAc is a late posttransitional event occurring in the Golgi apparatus (Nieman *et al.*, 1982; Johnson and Spear, 1983). Some studies have reported the transfer of GalNAc to the nascent polypeptide as a cotranslational event (Strous, 1979; Jokinen *et al.*, 1985). In the case of rat asialoglycoprotein ASGP-1, it was found that initiation of *O*-glycosylation occurred throughout the endomembrane system, from rough endoplasmic reticulum to the plasma membrane (Spielman *et al.*, 1988). The possibility remains that there are differences in the temporal aspects of the biosynthesis of *O*-glycosylated proteins in different tissues/cells and particularly in the synthesis of different types of molecules such as secreted mucins, membrane-associated *O*-glycosylated protein in normal cells, and *O*-glycosylated proteins in malignant cells. It is not surprising that the details of the synthesis of a mucin molecule having a very high carbohydrate content (70–80% by weight) as hundreds of separate saccharides are very different from those of the synthesis of glycoproteins having a low carbohydrate content and only a few *N*- and/or *O*-glycosyl units. Specifically the glycosylation of clusters of serine/threonine would present different problems than those of glycosylating lone (individual) Asn or Ser/Thr residues.

There is general agreement that, once the *O*-glycosyl chain is initiated, elongation and termination occurs rapidly as the molecule traverses the Golgi and *trans*-Golgi network to the secretory granules or cell surface. In pulse–chase

experiments generally it has not been possible to observe discrete intermediates probably because of the rapidity of the process. The chain elongation occurs by sequential addition of monosaccharides by highly specific glycosyltransferases and follows general patterns (Schachter and Brockhausen, 1992). Addition of Galβ→, GlcNAcβ→, or GalNAcα→ to positions 3 or 6 or both of the Ser/Thr-linked GalNAc will give rise to the core types 1 to 7 as illustrated in Figure 4. Schachter (1986) has postulated that the substitution of position 3 occurs before 6 (the 3 before 6 rule). Thus, core types 1 or 3 are first formed by introduction of Gal or GlcNAc, respectively, to position 3 of linkage GalNAc, and subsequent addition of GlcNAc to position 6 leads to core types 2 and 4. It was suggested that the same β1→6*N*-acetylglucosaminyltransferase catalyzes the synthesis of core 2 and 4 from core 1 and 3, respectively (Brockhausen *et al.*, 1985). In subsequent studies, it was found that the core 2 transferase activity but not core 4 activity was present in human leukocytes (Brockhausen *et al.*, 1991). Thus, the situation is apparently more complex, and recently it has been suggested that there may be three enzyme isoforms involved in the biosynthesis of the core 2 and 4 structures (Bierhuizen and Fukuda, 1993). Since the above activities cannot transfer GlcNAc to unsubstituted GalNAc, the synthesis of core 6 must be catalyzed by a different enzyme such as that reported by Yazawa *et al.* (1986). Cores 5 and 7 are synthesized probably by two separate GalNAc transferases which transfer GalNAcα to positions 3 and 6 of the linkage GalNAc, respectively.

In addition to the above transferases, α-*N*-acetylgalactosaminide α2→6 sialyltransferase can also act directly on GalNAc to yield the disaccharide SAα2→6GalNAc. This is a final product since it is not an acceptor for the enzymes responsible for synthesis of core 1 and 3 *in vitro* experiments. Even though the 3 before 6 rule would predict that core 6 and 7 structures are not substituted at positions 3 of GalNAc, such structures have been detected, albeit rarely, further illustrating the flexibility of the specificities of glycosyltransferases. Typically, core 2 and 4 are further galactosylated on the β1→6-linked GlcNAc before addition of terminal sugars. Chain termination occurs by the addition of sialic acid, fucose, and occasionally Gal, GalNAc, or GlcNAc all in α-anomeric linkage. Apparently, the α-linkage somehow prevents further addition of sugars and thereby inhibits chain elongation. Sulfation also occurs late in the biosynthesis of *O*-glycosyl chains and sulfate may be added to either the terminal or internal residues. In structures that have a long backbone, such as the repeating (Galβ1→4GlcNAcβ1→3) units, the synthesis of the repeating sequences takes place prior to sialylation or fucosylation. These repeating backbone structures are not unique to *O*-linked chains but also occur in Asn-linked chains and in some glycolipids (see Section 3.3).

An important point about *N*- or *O*-glycosylation is that the same protein can be differently glycosylated when expressed in different cell types. This was

clearly demonstrated for Thy-1 antigen from brain and thymus where it was established that the structures of the saccharides linked to the three Asn differed, but in a site-specific manner. Thus, in both cases $Asn_{23}$ contained only oligomannose structures and $Asn_{74}$ contained the complex sialylated saccharides (Parekh *et al.*, 1987). In the case of *O*-linked saccharides, striking differences in glycosylation are found in leukosialin saccharides of T cells, activated T cells, and neutrophils (Fukuda, 1991a). Resting T cells have the smaller core 1 structures whereas activated T cells and neutrophils have core 2 and elongated core 2 (polyactosamino) structures, respectively.

### 5.4. Biosynthesis of Polylactosaminoglycan Chains

Elongation of polylactosaminoglycans is initiated by addition of GlcNAc to the C-3 position of the galactose residue of the LacNAc group. The Galβ1→4-GlcNAc and Galβ1→Glc but not the Galβ1→3GlcNAc were found to be good acceptors for β-*N*-acetylglucosaminyltransferase (Piller and Cartron, 1983; Van den Eijnden *et al.*, 1988). Using desialylated $\alpha_1$-acid glycoprotein, fetuin and transferrin as acceptors for the elongation enzyme, GlcNAcTVII, have been shown to have branch specificity. Oligosaccharides with the Galβ1→4GlcNAc-β1→6(Galβ1→4GlcNAcβ1→2)Man group are more effective than those without this group (Van den Eijnden *et al.*, 1988). The branch specificity of the transferase is in accordance with the location of the polylactosamine structure on the 2,6-branched sugar chains of many glycoproteins. Recently, this glycosyltransferase has been purified to near homogeneity from calf serum (Kawashima *et al.*, 1993).

## 6. ONCOLOGY OF SIALOGLYCOPROTEINS

Studies in the 1960s demonstrated that sialidase treatment of cancer cells reduced their electrophoretic mobility (Abercrombie and Ambrose, 1962), caused substantial decreases in oncogenicity (Sanford, 1967), and increased immunogenicity in animals and protected them from subsequent challenges (Currie *et al.*, 1968). These observations sparked extensive investigations of the sialoglycoconjugates of malignant cells. Overall changes on the cell surface glycoconjugates were also evident from early studies on the differences in lectin binding and lectin-induced agglutination of cancer cells compared to normal cells. Soon many investigators documented that both quantitative and qualitative changes occur in cell surface sialoglycoproteins and gangliosides as a result of malignant transformation (Hakomori, 1989). These aberrations were considered to be the basis for the altered social behavior of cancer cells, such as uncontrolled

cell growth, altered cell adhesion, immunological resistance, invasiveness, and metastatic spread. Attenuation of the sugar chains, which is most widely observed in glycolipids (Hakomori, 1973), is also induced in the sugar chains of glycoproteins produced by transformed cells (Yamashita *et al.*, 1983). Since gangliosides are outside the scope of this chapter (see Chapters 6 and 7), only the alterations in the malignancy-associated sialoglycoproteins, both serum-type and mucin-type, will be discussed. For both classes of sialoglycoproteins, malignancy-associated changes which are of a general nature (i.e., overall, composite changes) and of a specific nature (changes on individual purified glycoproteins) have been documented (see review of Bhavanandan, 1991).

## 6.1. Malignancy-Associated Alterations in *N*-Glycosylated Sialoglycoproteins

One of the most prominent transformation-associated alterations in the sugar chains of glycoproteins is the increase of large Asn-linked sugar chains in the cell surface glycoproteins (see review by Kobata, 1989). This so-called Warren–Glick phenomenon was discovered by comparing the gel filtration patterns of the glycopeptides obtained by pronase digestion of metabolically labeled normal and malignant cell glycoproteins. Initially the increased molecular weight of the glycopeptides was considered to be related to a higher content of sialic acid in the sugar chains. However, subsequent studies revealed that the larger glycopeptides are not the result solely of increased sialylation but also of increased *N*-acetylglucosaminylation resulting in additional outer chains in the complex-type sugar chains. Detailed structural studies of the oligosaccharides isolated from BHK cells and polyoma- or RSV-transformed BHK cells proved that the increase of the GlcNAcβ1→6 branch attached to the Manα1→6Man arm of the trimannosyl cores is the structural basis for the Warren–Glick phenomenon (Yamashita *et al.*, 1984; Pierce and Arango, 1986). This was further confirmed by studies on a number of malignant cell lines transformed with different agents (Hiraizumi *et al.*, 1990; Santer *et al.*, 1984; Yousefi *et al.*, 1991).

Three transformed cell lines, MT1, MTPY, and MTAg, established from mouse NIH 3T3 cells by transfection with the SV40 or polyoma virus early gene segments (Segawa and Yamaguchi, 1986), showed no difference in anchorage-independent growth in soft agar gel, which is believed to be one of the characteristic features of transformed cells. However, when these cells were transplanted subcutaneously into athymic mice, they showed a marked difference in the tumorogenic potentials. MTAg and MTPY cells produced tumors 1 week and 2 months after inoculation, respectively, but MT1 and parental 3T3 cells did not form tumors even after 1 year (Figure 5). Furthermore, of the three cell lines, only MTAg showed pulmonary colonization when injected intravenously into athymic mouse. Therefore, the *in vitro* properties of transformed cells do not

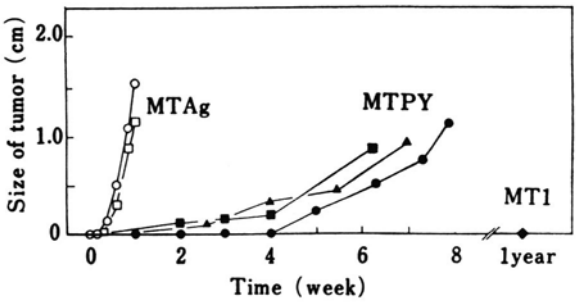



**FIGURE 5.** Tumor growth in athymic mice of NIH3T3 and its three transformed cell lines, MTAg, MTPY, and MT1. About $10^7$ cells were injected subcutaneously and tumor growth monitored weekly by measuring the mean geometric diameter (cm). MTAg (○, □), MTPY (●, ■, ▲), and MT1 (◆) cells were established by transfection with the polyoma middle T gene linked to the polyoma membrane attachment gene, SV40 large T gene linked to the polyoma membrane attachment gene and SV40 large gene, respectively (Segawa and Yamaguchi, 1986). While MTAg and MTPY cells formed tumors, MT1 and the parental NIH3T3 cells did not form tumors even after 1 year.

always reflect the *in vivo* properties. Structural analysis of the *N*-linked glycopeptides from these cells revealed that only about 20% of the glycopeptides from 3T3 and MT1 cells had 2,6-branched sugar chains compared to 31 and 39% of the glycopeptides from MTPY and MTAg cells, respectively (Asada *et al.*, 1992), indicating a proportionality between the increased 2,6-branching and tumorigenicity and/or metastatic potential. The increased expression of the 2,6-branched sugar chains is also correlated with tumor-forming activities of various other transformed cells (Hiraizumi *et al.*, 1990, 1992). The increased cell surface binding of fluorescein-labeled L-PHA, which specifically binds to the 2,6-branched sugar chains, was found to be correlated with the metastatic potentials of mouse mammary carcinoma cells, mouse lymphoma cells, transformed rat fibroblasts, and human breast cancers (Dennis *et al.*, 1987; Laferte and Dennis, 1989). Thus, the results of several independent studies establish that the Warren–Glick phenomenon caused by the increase of 2,6-branched sugar chains is actually related to the *in vivo* tumor-forming and metastatic potentials of transformed cells rather than transformation *per se*.

In accordance with the above structural studies, the specific activity of GlcNAcTV, which synthesizes the β-1,6 branch, has been shown to be elevated two- to three-fold in the transformed cells (Yamashita *et al.*, 1985; Pierce *et al.*, 1987; Yousefi *et al.*, 1991). In contrast, no change in the specific activities of other glycosyltransferases involved in the synthesis of the *N*-linked sugar chains was observed in transformed cells (Yamashita *et al.*, 1985). These findings confirm that the increase in the β-1,6-branched chains results from the increased specific activity of GlcNAcTV in transformed cells, although the molecular

mechanism causing the change in the GlcNAcTV activity in transformed cells remains to be elucidated. GlcNAcTV has been purified from rat kidney and human lung carcinoma, and its cDNA clone has been isolated from a rat (Shoreibah *et al.*, 1993) and human cDNA libraries (Saito *et al.*, 1994). The expression of the rat cDNA in COS-7 cells showed a significant increase of this enzyme activity (Shoreibah *et al.*, 1993). Using cell lines transfected with the GlcNAcTV cDNA, it will be interesting to determine whether changes in GlcNAcTV activity can directly influence cell adhesion and metastatic activity.

Several lectin-resistant cell lines have been shown to acquire reduced tumorigenic and/or metastatic potentials as compared to the parental malignant cells. The wheat germ agglutinin (WGA)-resistant L18 cell line established from metastatic B16 F10 mouse melanoma cells (Lin *et al.*, 1982) showed reduced tumor formation and metastasis compared to parental cells when injected subcutaneously into C57BL/6j mice (Table II, Furukawa and Bhavanandan, unpublished data). Structural analysis of the sugar chains of WGA-resistant MDW4 cells established from highly metastasizing mouse MDAY-D2 tumor cells showed that the mutant cells contained only ungalactosylated sugar chains and lacked the tetraantennary complex-type, polylactosamine, sugar chains which are the major sugar chains of the parental cells (Dennis *et al.*, 1986). These findings indicate the importance of the sialylated tetraantennary sugar chains with polylactosamine units in tumor metastasis. In contrast, the *N*-linked sugar chains obtained from the WGA-resistant cell line, Wa4-b1, and its parental B16 mouse melanoma F1 cell line showed that both cell lines contain almost similar amounts of high-mannose-type and bi-, tri-, and tetraantennary complex-type structures. The only difference was that the tri- and tetraantennary branches of the saccharides from the lectin mutant were not sialylated but heavily fucosylated (Finne *et al.*, 1980). Taken together, these results suggest that the expression of the β-1,6 branch itself is not sufficient for transformed cells to acquire a metastatic ability. Since low

**Table II**
**Tumorigenicity and Metastatic Potentials of B16 Mouse Melanoma F10 and Its WGA-Resistant L18 Cells[a]**

| | | B16 F10 cells | L18 cells |
|---|---|---|---|
| Tumorigenicity | 20 days[b] | 10/10, 10/10[c] | 6/10, 5/10 |
| | 30 days | All dead | 7/10, 6/10 |
| Metastatic potentials (lung, lymph node) | | 9/10, 9/10 | 2/10, 2/10 |

[a]Furukawa and Bhavanandan (unpublished results).
[b]Days after subcutaneous transplantation of $5 \times 10^4$ cells into C57BL/6 mouse.
[c]Number of animals that formed tumors out of a total of 10. Data from two experiments are shown.

sialylation in transformed cells has been shown to be associated with loss of metastasis (Yamamura *et al.*, 1991), the reduced sialylation caused by the fucosylation of the outer chain moieties in the mutant is considered to be a major factor in reducing the metastatic ability. Consistent with these observations, the different forms of the *N*-linked sugar chains have been shown to be involved in several recognition steps of metastasis. B16 mouse melanoma cells treated with tunicamycin, which blocks the oligosaccharide transfer to peptides, showed a reduced ability to adhere to vascular endothelial cells (Irimura *et al.*, 1981). Treatment of B16 mouse melanoma cell lines with swainsonine, which inhibits the processing of the *N*-linked sugar chains, resulted in inhibition of pulmonary colonization via reduced attachment to lung cells (Humphries *et al.*, 1986) and in *in vitro* invasion of basement membranes (Yagel *et al.*, 1989). Then, what is the role of the altered carbohydrate structures in transformed cells? Two possible roles are indicated by the following studies.

1. Polylactosaminoglycans may mediate cell adhesion. It is found that sialyl $Le^x$ and sialyl $Le^a$ determinants are expressed preferentially on the nonreducing ends of polylactosaminoglycan chains (Spooncer *et al.*, 1984) which are carried primarily by lysosome-associated membrane proteins, LAMP-1 and LAMP-2, respectively (Do *et al.*, 1990; Carlsson and Fukuda, 1992). Even though LAMPs are major constituents of lysosomal membranes, a small portion of these molecules are present on the cell surface (Lippincott-Schwartz and Fambrough, 1987; Carlsson and Fukuda, 1992). It was found that LAMP-1 is the major L-PHA-reactive glycoprotein in metastatic tumor cells (Chen *et al.*, 1988; Laferte and Dennis, 1989). In recent studies, analysis of the glycoproteins from MTAg cells revealed three major L-PHA-reactive bands one of which was tentatively identified as LAMP-1 (Asada, M., Segawa, K., Kobata, A., Endo, T., and Furukawa, K., unpublished). Tumor cells are believed to adhere to vascular endothelial cells by selectin-mediated interactions (Rice and Bevilacqua, 1989; Hession *et al.*, 1990) and the LAMPs are thought to provide some of the sialyl $Le^x$ and sialyl $Le^a$ ligands for this interaction. In fact, genetic manipulation of the level of LAMP-1 expression on colonic carcinoma cells clearly demonstrated that the increase of cell surface LAMP-1 is proportional to the extent of cell surface sialyl $Le^x$ and the binding to E-selectin (R. Sawada *et al.*, 1993). These studies indicate that there is a strong correlation between the metastatic potential of tumor cells and adhesion of these cells to endothelial cells mediated by the sialyl $Le^x$ determinants expressed on LAMP-1 of the tumor cell surface. Since GlcNAcTVII, which is involved in the synthesis of polylactosaminoglycans as described earlier, transfers GlcNAc effectively to the terminal galactose residue of the β-1,6 branch of the pentasaccharide Galβ1→4GlcNAcβ1→6(Galβ1→4GlcNAcβ1→2)Man (Van den Eijnden *et al.*, 1988), the significance of the increased expression of the 2,6-branched sugar chains in transformed cells may be to provide the necessary polylactosamine chains and hence selectin ligands.

2. The altered glycosylation may impair protein functions. In the case of the fibronectin receptor, it is known that the proper glycosylation of the receptor is important for binding to fibronectin. Akiyama *et al.* (1989) showed that human fibroblasts cultured in the presence of 1-deoxymannojirimycin, an inhibitor of Golgi α-mannosidase-I, does not bind well to fibronectin in spite of the presence of the receptor on the cell surface. The fibronectin receptor isolated from the drug-treated cells is also defective in the binding to immobilized fibronectin. These results indicate that the processed (mature) structures of the Asn-linked sugar chains are required for a functional fibronectin receptor. Analysis of the carbohydrate structures of the fibronectin receptor isolated from B16 mouse melanoma cells and the WGA-resistant cell line, which shows diminished binding to fibronectin and laminin (Tao and Johnson, 1982), revealed that the outer chains of the Asn-linked saccharides of the fibronectin receptor is important for binding to fibronectin (Kawano *et al.*, 1993). The parental cell, whose fibronectin receptor contained the sialylated tetraantennary structures, showed strong binding to fibronectin while the mutant cell expressing the receptor without these structures showed reduced binding. Further, the fibronectin receptor is one of the major L-PHA-reactive bands in MTAg cell glycoproteins, and the malignant MTAg cells but not the control 3T3 cells dissociate easily from the fibronectin substrate on treatment with anti-fibronectin receptor antibody (Asada, M., Segawa, K., Kobata, A., Endo, T., and Furukawa, K., unpublished). Therefore, the structural alterations occurring on the sugar chains of the fibronectin receptor may cause the transformed cells to adhere poorly, which is critical for the cells in the process of escaping from the primary tumor.

## 6.2. Malignancy-Associated Alterations in *O*-Glycosylated Sialoglycoproteins

The examination of the proteolytic fragments from normal and cancer cells provided evidence for the enrichment of higher-molecular-weight mucin-type sialoglycopeptides in a number of malignant cell lines. This was clearly demonstrated by studies on mouse and human melanoma cells; for example, pronase digestion of [$^3$H]glucosamine-labeled human melanoma cells and melanocytes followed by Bio-Gel P-10 gel filtration showed that five times more melanoma glycopeptides eluted at the void volume compared to the melanocyte glycopeptides (Bhavanandan *et al.*, 1977, 1981). Similarly increased production of mucin-type sialoglycopeptides by Morris hepatoma (Furukawa *et al.*, 1986), human breast cancer cells (Chandrasekaran and Davidson, 1979), and rat hepatomas AH66 and AH 130 cells (Funakoshi *et al.*, 1974; Nakada *et al.*, 1975) compared to their nonmalignant counterparts was demonstrated suggesting that this is also a common phenomenon associated with malignancy like the Warren–Glick phenomenon discussed above. Unfortunately, there were a number of

limitations to these (earlier) studies which made it difficult to assess the significance of the findings. For example, the normal cells such as hepatocytes or melanocytes used in some of these studies are not the most appropriate controls. Further, for studies on tissue culture cells it is difficult to determine whether the observed increase is malignancy-related or growth-related. More importantly, since the proteolytic fragments were a composite of many cellular glycoproteins whose identities are unknown, it is not possible to determine the functional implications of the changes. Additional evidence that the increased production of *O*-glycosylated protein was a general characteristic of all cancer cells came from subsequent immunochemical studies. The major oligosaccharides from the mucin-type glycopeptides of mouse and human melanomas were identical to those of glycophorin. Thus, polyclonal and monoclonal antiglycophorin antibodies were used to examine malignant and nonmalignant cells (Barsoum *et al.*, 1985; Barsoum and Bhavanandan, 1989). The presence of glycophorinlike epitopes was demonstrated, in varying concentrations, on a variety of cells including erythroleukemia, melanoma, and carcinomas of breast, cervix, larynx, nasopharynx, and colon (Barsoum *et al.*, 1984). However, these epitopes were not confined to malignant cells since cultured fibroblasts and HBL-100 (normal mammary) cells also gave positive immunofluorescences, albeit weak, suggesting quantitative rather than qualitative differences in the production of these glycoproteins.

The known changes in the *O*-glycosylated proteins associated with malignancy can be categorized into three types. The first are changes that occur in the peripheral sugars and backbone regions as a result of either incomplete synthesis or neosynthesis (Hakomori, 1984). This type of changes occurs typically in the lacto-series structures carrying blood group ABH, Lewis, Ii antigens and are thus common to both *N*- and *O*-linked glycoproteins as well as glycolipids. Examples of this type of change are the expression of sialyl $Le^a$ structures in gastrointestinal and pancreatic cancer-associated mucin glycoproteins (Magnani *et al.*, 1983) and the elongation of the backbone structures of $Le^x$ and $Le^Y$ antigens in malignant and premalignant colon cells (Kim *et al.*, 1986). The second type are changes that occur in the core region such as expression of the usually cryptic core structures as a result of blocked synthesis. The classic example of this is the incomplete synthesis and accumulation of GalNAcα→(Tn antigen) and Galβ1→3GalNAcα→(T antigen) structures on *O*-glycosylated proteins in adenocarcinomas (Springer, 1984). In addition to these, there are also increased levels of the disaccharide Neu5Ac2→6GalNAcα→(sialyl Tn antigen) in tumor-associated mucin glycoproteins (Johnson *et al.*, 1986). This can be explained by the conversion of some of the accumulated Tn antigen to sialyl Tn antigen by a specific α2→6 sialyltransferase. These Neu5Acα 2 →GalNAcα→ structures will accumulate since, as mentioned above, this disaccharide is a poor substrate

for the chain-elongating glycosyltransferases. The third type of change is the increased production of *O*-glycosylated proteins containing clusters of saccharides, and therefore resistant to proteolytic attack. This is illustrated by the above-discussed isolation of increased quantities of pronase-resistant mucin-type glycopeptides from cancer cells.

A great number of monoclonal antibodies have now been generated that distinguish between malignant cells and their nonmalignant counterparts based on qualitative (structural) and/or quantitative differences in cell surface *O*-glycosylated proteins. Some of the antibodies, such as MLS102, B72-3, and N19-9, are directed against carbohydrate epitopes on these proteins which have resulted as a consequence of the above-mentioned aberrations in glycosylation (Hakomori, 1989). The majority of these antibodies are, however, directed against peptide epitopes which are apparently cryptic on the *O*-glycosylated proteins of the nonmalignant cells. Even though there is no direct proof, it is believed that more of the peptide epitopes of mucin-type glycoproteins are exposed on malignant cell surfaces as a result of impaired glycosylation. Thus, it is possible that blocked synthesis leading to premature chain elongation would lead to (1) expression of short saccharide chains such as GalNAcα→ and Neu5Acα2→6GalNAcα→ epitopes, (2) exposed peptide epitopes which would otherwise have been covered by the larger saccharide structures, and (3) changed conformation of peptide regions creating new epitopes. Similarly, clustering of saccharides which are normally more evenly distributed on the core protein would also result in exposure of peptide epitopes on the malignancy associated *O*-glycosylated proteins and, in addition, create new multivalent saccharide epitopes. In fact, the monoclonal antibodies MLS102 and MLS128 with specificities toward malignant cells have been shown to be against clusters of GalNAcα→ and Neu5Acα 2→6GalNAc, respectively (Kurosaka *et al.*, 1988; Nakada *et al.*, 1993). However, one cannot at present entirely rule out the possibility that the peptide epitopes on malignant cells could also be generated by abnormal gene expression and/or mutations (Yonezawa *et al.*, 1991). It appears that some of the antibodies require both *O*-linked saccharides and peptide structures either because the epitope consists of both as may be the case with the Ca-1 antibody (Ashall *et al.*, 1982) or because glycosylation influences peptide epitope conformation as illustrated by the expression of blood group MN antigens and FD-6 antigens (Matsuura *et al.*, 1989). Currently, many of the monoclonal antibodies that recognize malignancy-associated *O*-glycosylated proteins are being exploited for the development of clinically useful diagnostic and therapeutic reagents for the management of cancer (Hilkens, 1988; Hilgers *et al.*, 1989; Hakomori, 1989).

Investigations of the allotransplantable ascites subline of TA3 mammary adenocarcinoma provided the first clear evidence of the association of one specific mucin-type glycoprotein with malignancy. A heavily *O*-glycosylated pro-

tein, named epiglycanin, was found to be present in large quantities on the surface of the allotransplantable TA3-Ha and TA3-MM sublines, but not on the non-allotransplantable TA3-St subline (Codington and Haavik, 1992). This molecule consists of a single polypeptide chain of about 1300 amino acid residues, to which about 550 short saccharide chains are attached via serine or threonine. Carbohydrate represents 75–80% of the glycoprotein of 500 kDa. About 80% of the saccharides in epiglycanin consist of monosaccharides (GalNAc) and disaccharides (Gal$\beta$1$\rightarrow$3GalNAc), the balance being short sialylated saccharides (Van den Eijnden *et al.*, 1979). Based on the mannose content of the purified glycoprotein, it is estimated that one or two Asn-linked saccharide chains may be present in the molecule. The amino acid composition is typical of that of *O*-glycosylated proteins, with Ser and Thr constituting about 48%, and Pro, Gly, Ala, Glu, and Asp accounting for a further 42%. This glycoprotein is produced by TA3-Ha cells only when grown in the ascites and not when cultured *in vitro* or grown as solid tumors. The *in vitro* cultured cells, however, regain their ability to produce epiglycanin when passaged in the ascites. The implication of this and the relationship of this phenomenon to the tumorigenicity of the cells are not fully understand. Epiglycanin is estimated to account for about 1% of the cell's dry weight and its high concentration on the cell surface (estimated to be about $4 \times 10^6$ molecules/cell) was demonstrated by lectin and antibody absorption experiments. Further, shadow-casting electron microscopy shows epitectin molecules on the surface of TA4-Ha and TA3-MM cells as long filaments (450–500 nm in length; Miller *et al.*, 1977). Epiglycanin has been shown to be shed from the cell surface since it is detected in the ascites fluid as well as in the serum of mice bearing the ascites cells. Intracellular epiglycanin is present in two types of exocytotic vesicles one of which consists of large multivesicular sacs with diameters up to 3 mm. It is suggested that epiglycanin can be secreted into the extracellular milieu via a new mode involving the multivesicular sacs (Watkins *et al.*, 1991). Epiglycanin has been proposed to protect the tumor cells from the host immune system by masking of the H-2 histocompatibility antigens (Codington and Frim, 1983) and by immunosuppression (Fung and Longenecker, 1991).

An *O*-glycosylated protein analogous to epiglyanin, and called ascites sialoglycoprotein (ASPG-1), has been identified on ascites 13762 rat mammary adenocarcinoma cells (Carraway and Spielman, 1986). Based on leucine labeling, ASGP-1 is estimated to constitute at least 0.5% of the total cell protein (Sherblom *et al.*, 1980a). Compared to epiglycanin, ASGP-1 appears to be heavily sialylated since sialidase treatment of 13762 cells substantially reduced ruthenium red and ferritin labeling (Sherblom *et al.*, 1980b). ASGP-1 is easy to isolate since centrifugation of the tumor cell membranes in a mixture of 4 M guanidine HCl and cesium chloride yields in a single step ASPG-1 of greater than 95% purity. The two different (MAT-B1 and MAT-C1) ascites sublines of 13762

cells produce ASGP-1 molecules with carbohydrate contents of 67 and 73%, respectively. Consequently, the molecules from the two sublines have different molecular size and density. The amino acid and carbohydrate compositions of ASGP-1 are very different from those of epiglycanin. In contrast to the small saccharides of epiglycanin, the saccharides of ASGP-1 are larger than tetrasaccharides and mostly sialylated and further, in the case of MAT-B1 species, are also sulfated (Sherblom and Carraway, 1980). It has been postulated that at the cell surface ASGP-1 occurs as a heterodimeric complex with a transmembrane sialoglycoprotein, termed ASGP-2. The ASGP-1 shed into the ascites fluid, apparently by a proteolytic cleavage, is free of ASGP-2. The two glycoproteins, ASGP-1 and ASGP-2, are generated from a single precursor protein. ASGP-2 is a 120-kDa sialoglycoprotein rich in Asn-linked saccharides and has two cysteine-rich domains with sequences related to the epidermal growth factor family of proteins (Sheng *et al.*, 1992). The resistance of the 13762 cells to lysis by natural killer cells and rat spleen lymphocytes was dependent on the expression of the ASGP-1/ASGP-2 complex on the cell surfaces. For example, cells treated with tunicamycin, which inhibits *N*-glycosylation of ASGP-2 and hence cell surface expression of ASGP-1, were significantly more susceptible to natural killer cell-mediated lysis (Bharathan *et al.*, 1990). Studies on a series of 13762 NK cultured cells selected for differing metastatic potential demonstrated a correlation with the presence of ASGP and metastasis (Steck and Nicolson, 1983).

Interest in malignancy-associated *O*-glycosylated proteins was greatly stimulated when a number of investigators produced monoclonal antibodies that were found to react with this family of glycoconjugates. Many of these antibodies were generated against human milk fat globule (HMFG) glycoproteins (Ceriani *et al.*, 1983; Burchell *et al.*, 1987), and others were obtained by immunizing with human tumor cells, cell membranes, or partially purified sialoglycoproteins (Colcher *et al.*, 1981; Ashall *et al.*, 1982; Kufe *et al.*, 1984). The sialoglycoproteins recognized by these antibodies have been called by different names including HMFG-1 antigen, PAS-O mucin, MAM-6, episialin, epithelial membrane antigen, DF-3 antigen, Ca antigen, epitectin, polymorphic urinary mucin (PUM), and polymorphic epithelial mucin (PEM). Cloning and sequencing studies have revealed that the core protein of these sialoglycoproteins is the product coded for by a gene now referred to as the MUC1 gene. This gene codes for a core protein that appears to be expressed also in a variety of normal human glandular epithelia (Peat *et al.*, 1992) and found in normal body fluids such as milk, sweat, and urine (Swallow *et al.*, 1986). Studies involving a large number of monoclonal antibodies specific for MUC1 glycoprotein reveal a strong preferential binding to cancer cells, indicating overexpression. The basis of the overexpression of the MUC1 epitopes by malignant cells is not fully understood, but it is believed that both increased synthesis, e.g., there is a more than tenfold increase of MUC1

mRNA in breast carcinoma (Zaretsky *et al.*, 1990), and increased exposure of peptide epitopes as a result of aberrant glycosylation (discussed elsewhere) are contributory factors. Based on the metabolic labeling experiments and the studies on epiglycanin and ASGP-1, it appears that in general an increased quantity of mucin-type glycoprotein, including the MUC1 product, may be produced by malignant cells compared to the counterpart nonmalignant cells. However, direct quantitative determinations of mucin glycoprotein in cancer cells and their normal counterparts by a colorimetric assay or other nonimmunological methods have not been performed (Bhavanandan, 1991). In normal epithelial cells, MUC1 glycoprotein is mainly present at the apical surface (Zotter *et al.*, 1988). In tumor cells this glycoprotein is spread out over the entire cell surface because of the loss of polarization (Hilkens *et al.*, 1992).

The human MUC1 gene encodes a core protein composed of variable numbers of 20-amino-acid tandem repeats, an *N*-terminal domain containing a signal sequence and a splice site yielding two alternative products, a C-terminal segment containing a potential transmembrane sequence and a cytoplasmic tail (Gendler *et al.*, 1990; Ligtenberg *et al.*, 1990; Williams *et al.*, 1990). It has been suggested that the cytoplasmic tail might be interacting specifically with cytoskeletal elements (Parry *et al.*, 1990; Spicer *et al.*, 1991). The extensive polymorphism of the human MUC1 gene (Swallow *et al.*, 1987b) is explained as resulting from variations in the number of tandem repeats. The tandem repeat consists of 25% Ser and Thr serving as potential *O*-glycosylation sites, and a further 25% is constituted by Pro. The nonrepetitive domains contain about 27% Ser and Thr, providing additional potential *O*-glycosylation sites. The nonrepetitive segment between the tandem repeats and the putative transmembrane sequence also contains five Asn in the consensus sequence for potential glycosylation.

The saccharide structures of the MUC1 family of glycoproteins are very poorly characterized. The currently available limited information suggests both tissue-dependent variations (Bardales *et al.*, 1989; Hull *et al.*, 1989) and malignancy-associated alteration in the saccharides of MUC1 glycoproteins (Hanisch *et al.*, 1990). It has been proposed that a shift from larger to smaller saccharides is responsible for exposing new peptide epitopes in MUC1 glycoprotein of cancer cells that are masked in nonmalignant cells (Burchell *et al.*, 1989); however, the evidence for this is very weak.

The MUC1 glycoprotein produced by the human laryngeal carcinoma (H.Ep.2) cell line and named CA antigen or epitectin has been investigated in some detail (Ashall *et al.*, 1982; Bramwell *et al.*, 1983; Bardales *et al.*, 1989; Qin and Bhavanandan, unpublished results). Epitectin was originally discovered after noting the ability of the Ca1 monoclonal antibody recognizing it to distinguish between malignant and nonmalignant segregants of hybrid cells (Ashall *et al.*, 1982). Thus, in the pair of CGL1 and CGL3 cells derived from the fusion of

human cervical carcinoma cells (HeLa cells) and human diploid fibroblasts (Stanbridge *et al.*, 1982), the ability to produce epitectin is linked to the ability of the cells to grow progressively in vivo in nude mice. Additional evidence for the link between epitectin production and tumorigenicity comes from observations on the effect of long-term culture of H.Ep.2 cells. After 80 or more generations in culture the production of epitectin by H.Ep.2 cells was significantly reduced and this was accompanied by a decrease in the ability of these cells to produce tumors in nude mice (DiIulio *et al.*, 1994). Gel filtration of epitectin purified by affinity chromatography shows very heterogeneous elution profiles typical of epithelial mucins (Bardales *et al.*, 1989). Mucin-type characteristics are also evident from the heterogeneous populations of flexible strands of epitectin molecules visualized by electron microscopy (Bramwell *et al.*, 1986). However, on SDS-PAGE the epitectin preparations are surprisingly (compared to secreted epithelial mucins) homogeneous showing two distinct, reasonably sharp bands with estimated molecular masses of 350 and 390 kDa. A buoyant density of about 1.4 g/ml for this glycoprotein is suggestive of a carbohydrate content of about 50% by weight. The bulk of the carbohydrate is present as di-, tri-, and tetrasaccharides of the general formula (NeuAcα) $_{0,1,2}$→[Galβ1→3GalNAc]α→. Other minor saccharides detected include GalNAcα→(Tn antigen), SAα2→6-GalNAcα→ (sialyl Tn antigen), and the hexasaccharide NeuAc→Gal(NeuAc→Gal→GlcNAc)→GalNAcα→. A small amount of mannose is present suggesting the presence of at least one Asn-linked oligosaccharide. Proteolytic fragmentation followed by HPLC and lectin affinity chromatography revealed the presence of peptide regions with saccharide clustering and regions free of glycosylation (Bardales, DiIulio, and Bhavanandan, unpublished results). Sequential extractions with ionic buffers and detergents indicate that epitectin is an intrinsic membrane glycoprotein which is not released by treatment with phospholipase C. It is, therefore, most likely held by a transmembrane peptide segment as suggested by the deduced amino acid sequence.

It is now abundantly clear that *O*-glycosylated proteins are intimately associated with cell surfaces and that in malignant cells there are both quantitative and qualitative changes in these molecules. However, information concerning the function of these molecules in normal and malignant cells is sparse (see Carraway *et al.*, 1992). It is generally accepted that the secretory mucins, which form viscoelastic gels on epithelial surfaces that are in direct contact with the environment, provide lubrication and protection. Thus, the respiratory tract mucous secretions serve to trap and clear inhaled pathogens (virus, bacteria), particulate pollutants (such as pollen or dust), and corrosive agents (such as sulfur dioxide) via the mucociliary transport system, thereby maintaining the sterility of the lungs (Kaliner, 1991). Similarly, vaginal/cervical mucins are important for maintaining the sterility of the uterus. The gastrointestinal mucus impedes the diffusion of $H^+$ ions into the epithelium and prevents digestion of the epithelial

cells by acids as well as by enzymes (Allen, 1983; Bhaskar *et al.*, 1992). In lower forms of animals, such as earthworms, slugs, frogs, and fish, secretions that cover the outer body surfaces function as mechanical protectant, hydrodynamic lubricant, and probably also as antiparasitic and antibacterial agents (see Ingram, 1980; Bevins and Zasloff, 1990). In serving these functions, the secretory mucins play two distinct roles: that of a mechanical protectant/lubricant gel and that of specific receptors for pathogenic organisms. The ability of the mucins to form viscous solutions or a gel is clearly dependent on the saccharides, but a saccharide–protein glycoconjugate is not essential since a pure polysaccharide, such as hyaluronic acid, has this property. In mucins, a core protein molecule is heavily substituted with small saccharides and thus resembles polysaccharides in an overall sense. Examples of some such "simple" mucins are armadillo and sheep submaxillary mucins in which the major portion of the carbohydrates are *O*-linked GalNAc and SAα2→6GalNAc, respectively. It has been demonstrated that attachment of GalNAc residues leads to stiffening of the core polypeptide chain resulting in highly hydrophilic, random-coil molecules (Shogren *et al.*, 1989). Addition of an extra sugar, particularly a charged sugar such as sialic acid, would increase hydrophilicity and gel-forming ability (Gottschalk and Thomas, 1961; Maeji *et al.*, 1987). Thus, it seems unlikely that larger oligosaccharides and the enormous structural diversity of the saccharides noted in the majority of mucins are essential for the protective/lubricative role of mucins. More than 60 separate saccharides were identifiable in human tracheobronchial mucins and undoubtedly more will be found by improvements in techniques of fractionation (Woodward, Ringler, Davidson, and Bhavanandan, unpublished results). An attractive explanation for this complexity is that these different saccharide structures are needed as specific receptors for capturing and eliminating pathogens (virus, bacteria, parasites). Varki (1993) in a recent review refers to this as the "decoy" function of oligosaccharides of mucins, as opposed to the "traitorous" function of mostly glycolipid-based saccharides on cell surfaces. There is no clear evidence that gel-forming secretory mucins are produced by malignant cells. All available evidence indicates that malignancy-associated *O*-glycosylated proteins are primarily intrinsic membrane molecules, even though some portions are shed/secreted into the extracellular milieu. The observation that epiglycanin is packaged in multivesicular sacs is important and should be pursued to determine whether it is actively secreted and if this glycoprotein has the ability to yield viscous solutions and gels.

It is speculated that the membrane-associated/anchored mucin-type glycoproteins of malignant cells also play protective roles. Some of these were mentioned above with respect to possible functions of epiglycanin and ASGP-1. It has been suggested that mucin-type glycoproteins associated with the surface of cancer cells may serve to shield these cells from an abnormal extracellular environment such as low pH, altered osmolarity and hydrolytic enzymes that would otherwise be destructive (Bramwell *et al.*, 1983). The high glycolytic

activity of cancer cells is likely to result in low pH within the tumor, and cell surface mucin-type glycoprotein could play a role similar to that of gastric mucin (Bhaskar *et al.*, 1992). There may be a common function for the intrinsic mucin-type glycoproteins associated with erythrocytes, leukocytes, platelets, and a wide variety of noncancerous cells. An understanding of any such common or of specific functions of glycophorin, leukosialin, CD42b, GYCAM-1, CD34, and the range of MUC1 glycoprotein-positive normal cells (Zotter *et al.*, 1988) would greatly accelerate the elucidation of the role of this class of glycoconjugate associated with the surface of cancer cells. On the cell surface, these extended rigid molecules are likely to project high above the plasma membrane into the lumen and thereby shield the surface and regulate the interaction with other cells, membranes, and soluble components. For MUC1 glycoprotein a variety of effects have been noted in *in vitro* experiments. Shimizu *et al.* (1990) found that the human milk glycoprotein (HMFG) was able to markedly reduce the proliferation rate of BALB/c 3T3 cells in culture in a concentration-dependent and reversible manner, suggesting a role in the regulation of cell growth. Ligtenberg *et al.* (1992) have proposed that MUC1 glycoproteins are antiadhesion molecules which function by masking adhesion molecules that are present on the cell membrane, as well as by charge repulsion. When normal mammary epithelial cells and melanoma cells which do not express exogenous MUC1 were transfected with cDNA encoding MUC1 glycoprotein, the resulting transfectants (expressing MUC1 glycoprotein at levels similar to carcinoma cells) did not aggregate as efficiently as the control (nontransfected) cells. This could be one of the factors that influence the metastatic potential of tumor cells. MUC1 glycoprotein has been implicated in the protection of tumor cells from the immune system by different mechanisms. It could simply cause steric hindrance and interfere with the recognition of surface antigens as proposed to be the case with epiglycanin. Purified MUC1 glycoprotein has been shown to inhibit the cytotoxic action of eosinophils toward target cells (Hayes *et al.*, 1990). Cytotoxic T lymphocytes derived from patients with breast cancer were found to be able to recognize MUC1 glycoprotein as their target in a major histocompatibility complex (MHC)-unrestricted fashion (Jerome *et al.*, 1991). It is suggested that the highly repetitive nature of the MUC1 epitopes allows cross-linking of T-cell receptors directly, accounting for the lack of MHC restriction. In recent studies, van de Wiel-van Kemenade *et al.* (1993) transfected MUC1-minus melanoma cells with MUC1 cDNA and found that the transfected cells were significantly less susceptible to lysis by cytotoxic lymphocytes compared to the parental cells. Finally, the MUC1 glycoprotein shed from the cells into the circulation may also act as blocking factors and inhibit lysis of MUC1-expressing cells by the cytotoxic T cells as noted in studies with soluble MUC1 glycoprotein isolated from milk (Barnd *et al.*, 1989).

The metastatic process is very complex and involves numerous steps including detachment of cells from the primary tumor, motility, extravasation,

escape from immunosurveillance, and adhesion/invasion of target organs. Various cell surface saccharide structures and whole glycoconjugate molecules are thought to be important in these steps. Numerous investigators have examined the correlation between alterations in the sialylation of cell surface glycoconjugates and metastatic potential of tumor cells. While some studies indicated that highly metastatic cells had elevated levels of sialic acid (Yogeswaran and Tao, 1980; Altevogt *et al.*, 1983) others did not find significant differences in the cell surface sialic acid of metastatic variants (Passaniti and Hart, 1988). The biphasic effect of cell surface sialic acids on cancer cell adhesiveness recently reported by T. Sawada *et al.* (1993) could be an explanation for some of the conflicting findings on the correlation between sialic acid levels and invasiveness and metastasis of tumor cells. The increased levels of GlcNAcβ→6Manα→6Manβ→ branches in Asn-linked oligosaccharides and of GlcNAcβ→6 branches in Se/Thr-linked oligosaccharides have been found to be correlated with tumor progression and increased metastatic potential of tumor cells as discussed above (Dennis *et al.*, 1987; Laferte and Dennis, 1989). Studies on T24H-ras-infected fibroblasts and SP1 mammary carcinoma cells have shown increased levels of the glycosyltransferases that are responsible for the synthesis of the above structures (Yousefi *et al.*, 1991). Recently, the involvement of cancer cell surface *O*-glycosylated proteins in metastasis has received special attention because of the belief that these molecules could have multiple roles in the complex process of metastatic spread of cancer. For example, *O*-glycosylated proteins such as the MUC1 glycoprotein when present at some optimum level could promote cell–cell adherence necessary for the growth of the tumor, but at high levels it could cause poor adhesion and escape into the circulation. Once in the circulation, the cells with high levels of *O*-glycosylated proteins would be in an advantageous position to escape destruction by the immune system. The ability of these glycoproteins to suppress NK cell cytotoxicity could be one of the factors in this process of evading the immune system. In fact, a number of investigations have shown that the production of high levels of mucin glycoproteins by tumors is correlated to poor prognosis in cancer patients (Symond and Vickery, 1976; Ater *et al.*, 1984). In colorectal cancer, the presence of a high level of sialyl Tn antigen on *O*-glycosylated proteins of tumor cells was found to be a predictor of poor prognosis suggesting that in addition to quantitative differences structural alterations in the molecule are also important (Itzkowitz *et al.*, 1990). A number of investigations have indicated that high-molecular-weight *O*-glycosylated proteins are differentially expressed on cultured metastatic human cancer cells (Irimura *et al.*, 1988; Hoff *et al.*, 1989; Bresalier *et al.*, 1991). Other studies involving a large number of human colorectal carcinoma specimens have also demonstrated a strong association with increased expression of this class of glycoproteins (in this case, MUC1 glycoprotein) and progression to advanced

stages and metastasis of human colorectal carcinoma (Nakamori *et al.*, 1994). In this study, since aberrant transcription of MUC1 core protein did not appear to be the reason for the increased expression, altered processing or increased stability of the mature MUC1 glycoprotein is thought to be responsible for the higher levels. Altered processing of MUC1, such as premature termination of saccharide chains, would explain both the increased exposure of peptide epitopes (Zotter *et al.*, 1988) and the increase in Tn and sialyl Tn antigens detected by various antibodies (Springer, 1984; Itzkowitz *et al.*, 1990). These are not the only changes in mucin glycoprotein-associated carbohydrate antigens. Thus, the levels of sialyl-dimeric $Le^x$ antigen expressed on *O*-glycosylated proteins in colon carcinoma cells that metastasized to the liver were clearly higher than those on the primary tumor cells (Hoff *et al.*, 1989; Matsushita *et al.*, 1991). These findings are very important since the sialyl $Le^x$ structure is implicated to be a major ligand for the lectinlike adhesion molecules, E- and P-selectins (Irimura *et al.*, 1993). Interactions between selectins and ligands such as sialyl $Le^x$ saccharide on the tumor cell surface could mediate adhesion to endothelial cells and to platelets involved in the process of extravasation (Irimura *et al.*, 1993). Sawada *et al.* (1994) found that the efficiency of the E-selectin-mediated binding of colon cancer cells to endothelial cells correlated with the metastatic potential of the cancer cells. The E-selectin-mediated binding of the tumor cells could be inhibited by glycoproteins such as leukosialin and LAMP-1 which contain sialyl $Le^x$ structures, but not by other glycoproteins. The colon carcinoma-associated overexpression of sialyl $Le^x$ was found to be the result of altered glycosylation of MUC1 core protein (Hanski *et al.*, 1993). Thus, evidence is accumulating for an intimate involvement of *O*-glycosylated sialoglycoproteins in general and MUC1 sialoglycoprotein in particular with the process of tumorigenicity and metastasis. However, at present very little is known concerning the factors that control the expression of this class of glycoconjugates at the level of the gene as well as the processing and glycosylation of mucin-type glycoproteins in normal and tumor cells. Irimura and co-workers (Irimura *et al.*, 1990; Dohi *et al.*, 1993; Shirotani *et al.*, 1994) have identified a soluble factor, named mucomodulin, in human colon tissues which stimulates the production of MUC1 glycoprotein by upregulation of the mRNA levels. Purification and characterization of mucomodulin and understanding of its role in the regulation of MUC1 gene expression would greatly enhance our understanding of the role of the high-molecular-weight *O*-glycosylated proteins in malignancy.

ACKNOWLEDGMENTS. The investigations of one of us (V.P.B.) in this field have been supported by grants from the USPHS, most recently by grant CA38797 awarded by the National Cancer Institute. We thank Ms. Esha Bhavanandan and Ms. Filomena Cramer for their secretarial assistance.

## REFERENCES

Abeijon, C., and Hirschberg, C. B., 1987, Subcellular site of synthesis of the N-acetylgalactosamine ($\alpha$1→0) serine (or threonine) linkage in rat liver, *J. Biol. Chem.* **262:**4153–4159.

Abercrombie, M., and Ambrose, E. J., 1962, The surface properties of cancer cells: A review, *Cancer Res.* **22:**525–548.

Akiyama, S. K., Yamada, S. S., and Yamada, K. M., 1989, Analysis of the role of glycosylation of the human fibronectin receptor, *J. Biol. Chem.* **264:**18011–18018.

Allen, A., 1983, Mucus—A protective secretion of complexity, *Trends Biochem. Sci.* **8:**169–173.

Altevogt, P., Fogel, M., Cheingsong-Popov, R., Dennis, J. W., Robinson, P., and Schirrmacher, V., 1983, Different patterns of lectin binding and cell surface sialylation detected on related high- and low-metastatic tumor lines, *Cancer Res.* **43:**5138–5144.

Asada, M., Furukawa, K., Kantor, C., Gahmberg, C. G., and Kobata, A., 1991, Structural study of the sugar chains of human leukocyte cell adhesion molecules CD11/CD18, *Biochemistry* **30:**1561–1571.

Asada, M., Furukawa, K., Segawa, K., and Kobata, A., 1992, Biological significance of altered glycosylation of the cell surface glycoproteins in the transformed cells, *Cell Struct. Funct.* **17:**477 (Abstract).

Ashall, F., Bramwall, M. E., and Harris, H., 1982, A new marker for human cancer cells. 1. The Ca antigen and the Ca1 antibody, *Lancet* **7:**1–6.

Ater, J. L., Gooch, W. M., Bybee, B. L., and O'Brien, R. T., 1984, Poor prognosis for mucin-producing Wilms' tumor, *Cancer* **53:**319–323.

Aubert, J., Biserte, G., and Loucheux-Lefebvre, M. H., 1976, Carbohydrate–peptide linkage in glycoproteins, *Arch. Biochem. Biophys.* **175:**410–418.

Bardales, R., Bhavanandan, V. P., Wiseman, G., and Bramwell, M. E., 1989, Purification and characterization of the epitectin from human laryngeal carcinoma cells, *J. Biol. Chem.* **264:**1980–1987.

Barnd, D. L., Lan, M. S., Metzgar, R. S., and Finn, O. K., 1989, Specific major histocompatibility complex-unrestricted recognition of tumor-associated mucins by human cytotoxic T cells, *Proc. Natl. Acad. Sci. USA* **86:**7159–7163.

Barsoum, A. L., and Bhavanandan, V. P., 1989, Detection of glycophorin A-like glycoproteins on the surface of cultured human cells, *Int. J. Biochem.* **21:**635–654.

Barsoum, A. L., Czuczman, M., Bhavanandan, V. P., and Davidson, E. A., 1984, Epitopes immunologically related to glycophorin A on human malignant and nonmalignant cells in culture, *Int. J. Cancer* **34:**789–795.

Barsoum, A. L., Bhavanandan, V. P., and Davidson, E. A., 1985, Monoclonal antibodies to cyanogen bromide fragments of glycophorin A, *Mol. Immunol.* **22:**361–367.

Bevilacqua, M. P., and Nelson, R. M., 1993, Selectins, *J. Clin. Invest.* **91:**379–287.

Bevins, C. L., and Zasloff, M. A., 1990, Peptides from frog skin, *Annu. Rev. Biochem.* **59:**395–414.

Beyer, T. A., and Hill, R. L., 1982, Glycosylation pathways in the biosynthesis of nonreducing terminal sequences in oligosaccharides of glycoproteins, in: *The Glycoconjugates* III (M. I. Horowitz, ed.), Academic Press, New York, pp. 25–45.

Bharathan, S., Moriarty, J., Moody, C. E., and Sherblom, A. P., 1990, Effect of tunicamycin on sialomucin and natural killer susceptibility of rat mammary tumor ascites cells. *Cancer Res.* **50:**5250–5256.

Bhargava, A. K., Woitach, J. J., Davidson, E. A., and Bhavanandan, V. P., 1990, Cloning and cDNA sequence of a bovine submaxillary gland mucin-like protein containing two distinct domains, *Proc. Natl. Acad. Sci. USA* **87:**6798–6802.

Bhaskar, K. R., Garik, P., Turner, B. S., Bradley, J. D., Bansil, R., Stanley, H. E., and Lamont, J. T., 1992, Viscous fingering of HCl through gastric mucin, *Nature* **360:**458–462.

Bhavanandan, V. P., 1991, Cancer-associated mucins and mucin-type glycoproteins, *Glycobiology* **1:**493–503.

Bhavanandan, V. P., and Hegarty, J. D., 1987, Identification of the mucin core protein by cell-free translation of messenger RNA from bovine submaxillary glands, *J. Biol. Chem.* **262:**5913–5917.

Bhavanandan, V. P., and Katlic, A. W., 1979, The interaction of wheat germ agglutinin with sialoglycoproteins; The role of sialic acid, *J. Biol. Chem.* **254:**4000–4008.

Bhavanandan, V. P., and Meyer, K., 1966, Mucopolysaccharides: N-acetyl glucosamine- and galactose-6 sulfates from keratosulfate, *Science* **151:**1404–1405.

Bhavanandan, V. P., Umemoto, J., Banks, J. R., and Davidson, E. A., 1977, Isolation and partial characterization of sialoglycopeptides produced by a murine melanoma, *Biochemistry* **16:**4426–4437.

Bhavanandan, V. P., Katlic, A. W., Banks, J., Kemper, J. G., and Davidson, E. A., 1981, Partial characterization of sialoglycopeptides produced by cultured human melanoma cells and melanocytes, *Biochemistry* **20:**5586–5594.

Bierhuizen, M.F.A., and Fukuda, M., 1993, β-1,6-N-acetylglucosaminyl transferases: Enzymes critically involved in oligosaccharide branching, *Trends in Glycoscience and Glycotechnology* **6:**17–28.

Bobek, L. A., Tsai, H., Biesbrock, A. R., and Levine, M. J., 1993, Molecular cloning, sequence and specificity of expression of the gene encoding the low molecular weight human salivary mucin (Muc7), *J. Biol. Chem.* **268:**20563–20569.

Bramwell, M. E., Bhavanandan, V. P., Wiseman, G., and Harris, H., 1983, Structure and function of the Ca antigen, *Br. J. Cancer* **48:**177–183.

Bramwell, M. E., Wiseman, G., and Shotton, D. M., 1986, Electron-microscopic studies of the Ca antigen, epitectin, *J. Cell Sci.* **86:**249–261.

Bresalier, R. S., Niv, Y., Byrd, J. C., Duh, Q., Toribara, N. W., Rockwell, R. W., Dahiya, R., and Kim, Y. S., 1991, Mucin production by human colonic carcinoma cells correlates with their metastatic potential in animal models of colon cancer metastasis, *J. Clin. Invest.* **87:**1037–1045.

Briand, J. P., Andrews, S. P., Cahill, E., Conway, N. A., and Young, J. D., 1981, Investigation of the requirements for O-glycosylation by bovine submaxillary gland UDP-N-acetylgalactosamine: polypeptide N-acetylgalactosamine transferase using synthetic peptide substrates, *J. Biol. Chem.* **256:**12205–12207.

Brockhausen, I., Matta, K. L., Orr, J., and Schachter, H., 1985, Mucin synthesis. UDP-GlcNAc: GalNAc-Rβ 3-N-acetylglucosaminyltransferase and UDP-GlcNAc:GlcNAcβ1→3GalNAc-R (GlcNAc to GalNAc) β6-N-acetyl glucosaminyltransferase from pig and rat colon mucosa, *Biochemistry* **24:**1866–1860.

Brockhausen, I., Kuhns, W., Schachter, H., Matta, K. L., Sutherland, D. R., and Baker, M. A., 1991, Biosynthesis of O-glycans in leukocytes from normal donors and from patients with leukemia: Increase in O-glycan core 2 UDP-GlcNAc: Galβ3→GalNAcR (GlcNAc to GalNAc) β(1→6)-N-acetyl glucosaminyltransferase in leukemic cells, *Cancer Res.* **51:**1257–1263.

Burchell, J., Gendler, S., Taylor-Papadimitriou, J., Girling, A., Lewis, A., Millis, R., and Lamport, D., 1987, Development and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of the human milk mucin, *Cancer Res.* **47:**5476–5482.

Burchell, J., Papadimitriou, J. T., Boshell, M., Gendler, S., and Duhig, T., 1989, A short sequence, within the amino acid tandem repeat of a cancer-associated mucin, contains immunodominant epitopes, *Int. J. Cancer* **44:**691–696.

Carlsson, S. R., and Fukuda, M., 1992, The lysosomal membrane glycoprotein lamp-1 is transported to lysosomes by two alternative pathways, *Arch. Biochem. Biophys.* **296:**630–639.

Carraway, K. L., and Spielman, J., 1986, Structural and functional aspects of tumor cell sialomucins, *Mol. Cell. Biochem.* **72:**109–120.

Carraway, K. L., Fregien, N., Carraway, K. L., III, and Carraway, C.A.C., 1992, Tumor sialomucin complexes as tumor antigens and modulators of cellular interactions and proliferation, *J. Cell Sci.* **103:**299–307.

Cartron, J. P., and Rahuel, C., 1992, Human erythrocyte glycophorins: Protein and gene structure analysis, *Transfus. Med. Rev.* **11:**63–92.

Ceriani, R. L., Peterson, J. A., Leo, J. Y., Moncada, R., and Blank, E. W., 1983, Characterization of cell surface antigens of human mammary epithelial cells with monoclonal antibodies prepared against human milk fat globule, *Somat. Cell Genet.* **9:**415–427.

Chai, W., Hounsell, E. F., Cashmore, G. C., Rosaqnkiewicz, J. R., Bauer, C. J., Fenney, J., Feizi, T., and Lawson, A. M., 1992, Neutral oligosaccharides of bovine submaxxilary mucin, *Eur. J. Biochem.* **203:**257–268.

Chandrasekaran, E. V., and Davidson, E. A., 1979, Sialoglycoproteins of human mammary cells: Partial characterization of sialoglycopeptides, *Biochemistry* **18:**5615–5620.

Chen, J. W., Cha, Y., Yuksel, K. U., Gracy, R. W., and August, J. T., 1988, Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycoprotein mouse LAMP-1, *J. Biol. Chem.* **263:**8754–8758.

Codington, J. F., and Frim, D. M., 1983, Cell Surface macromolecular and morphological changes related to allotransplantability in the TA3 tumor, *Biomembranes* **11:**207–258.

Codington, J. F., and Haavik, S., 1992, Epiglycanin—A carcinoma-specific mucin-type glycoprotein of the mouse TA3 tumor, *Glycobiology* **2:**173–180.

Colcher, D., Horan, H. P., Nuti, M., and Schlom, J., 1981, A spectrum of monoclonal antibodies reactive with human mammary tumor cells, *Proc. Natl. Acad. Sci. USA* **78:**3199–3203.

Crepin, M., Porchet, N., Aubert, J. P., and Degand, P., 1990, Diversity of the peptide moiety of human airway mucins, *Biorheology* **27:**471–484.

Cummings, R. D., Trowbridge, I. S., and Kornfeld, S., 1982, A mouse lymphoma cell line resistant to the leukoagglutinating mannoside-β1,6 N-acetylglucosaminyltransferase, *J. Biol. Chem.* **257:**13421–13427.

Currie, G. A., Van Doorninck, W., and Bagshawe, K. D., 1968, Effect of neuraminidase on the immunogenicity of early mouse tropoblast, *Nature* **219:**191–192.

Dekker, J., and Strous, G. J., 1990, Covalent oligomerization of rat gastric mucin occurs in the rough endoplasmic reticulum, is N-glycosylation dependent, and precedes initial O-glycosylation, *J. Biol. Chem.* **265:**18116–18122.

Dekker, J., Aelmans, P. H., and Strous, G. J., 1991, The oligomeric structure of rat and human gastric mucins, *Biochem. J.* **277:**423–427.

Dennis, J. W., Laferte, S., Fukuda, M., Dell, A., and Carver, J. P., 1986, Asn-linked oligosaccharides in lectin-resistant tumor-cell mutants with varying metastatic potential, *Eur. J. Biochem.* **161:**359–373.

Dennis, J. W., Laferte, S., Waghorne, M. L., Breitman, M. L., and Kerbek, R. S., 1987, β1→6 branching of Asn-linked oligosaccharides is directly associated with metastasis, *Science* **236:**582–585.

DiIulio, N. A., Yamakami, K., Washington, S., and Bhavanandan, V. P., 1994, Effect of long-term culture of a human laryngeal carcinoma cell line on epitectin production and tumorigenicity in athymic mice, *Glycosylation Dis.* **1:**21–30.

Do, K.-Y., Smith, D. F., and Cummings, R. D., 1990, LAMP-1 in Cho cells is a primary carrier of poly-N-acetyllactosamine chains and is bound preferentially by a mammalian S-type lectin, *Biochem. Biophys. Res. Commun.* **173:**1123–1128.

Dohi, D. F., Sutton, R. C., Frazier, M. L., Nakamori, S., McIsaac, A. M., and Irimura, T., 1993, Regulation of sialomucin production in colon carcinoma cells, *J. Biol. Chem.* **268:**10133–10138.

Dufosse, J., Porchet, N., Audie, J.-P., Duperat, V. G., Laine, A., Van-Seuningen, I., Marrakchi, S.,

Degand, P., and Aubert, J.-P., 1993, Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic alternating domains in human mucin peptides mapped to 11p15, *Biochem. J.* **293:**329–337.

Eckhardt, A. E., Timpte, C. S., Abernathy, J. L., Zhao, Y., and Hill, R. L., 1991, Porcine submaxillary mucin contains a cystine-rich, carboxylterminal domain in addition to a highly repetitive, glycosylated domain, *J. Biol. Chem.* **266:**9678–9687.

Endo, T., Ohbayashi, H., Hayashi, Y., Ikehara, Y., Kochibe, N., and Kobata, A., 1988, Structural study on the carbohydrate moiety of human placental alkaline phosphatase, *J. Biochem.* **103:**182–187.

Fiete, D., Srivastava, V., Hindsgaul, O., and Baenziger, J. U., 1991, A hepatic reticuloendothelial cell receptor specific for $SO_4 \rightarrow$GalNAcβ1,4 GlcNAcβ1,2 Manα that mediates rapid clearance of lutropin, *Cell* **67:**1103–1110.

Finne, J., Tao, T. W., and Burger, M. M., 1980, Carbohydrate changes in glycoproteins of a poorly metastasizing wheat germ agglutinin-resistant melanoma clone, *Carbohydr. Res.* **40:**2580–2587.

Fukuda, M., 1991a, Leukosialin, a major O-glycan-containing sialoglycoprotein defining leukocyte differentiation and malignancy, *Glycobiology* **1:**347–356.

Funakoshi, I., Nakada, H., and Yamashina, I., 1974, The isolation and characterization of glycopeptides from plasma membranes of an ascites hepatoma, AH 66, *J. Biochem.* **76:**319–333.

Fung, P.Y.S., and Longenecker, B. M., 1991, Specific immunosuppressive activity of epiglycanin, a mucin-like glycoprotein secreted by a murine mammary adenocarcinoma (TA3-Ha), *Cancer Res.* **51:**1170–1176.

Furukawa, K., and Kobata, A., 1992, Protein glycosylation, *Curr. Opin. Biotechnol.* **3:**554–559.

Furukawa, K., Minor, J. E., Hegarty, J. D., and Bhavanandan, V. P., 1986, Interactions of sialoglycoproteins with wheat germ agglutinin-Sepharose of varying ratio of lectin to sepharose: Use for the purification of mucin glycoproteins from membrane extracts, *J. Biol. Chem.* **261:**7755–7761.

Galili, U., Shohet, S. B., Kobrin, E., Stults, C. L., and Macher, B. A., 1988, Man, apes, and old world monkeys differ from other mammals in the expression of alpha-galactosyl epitopes on nucleated cells, *J. Biol. Chem.* **263:**17755–17762.

Gendler, S. J., Lancaster, C. A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E., and Wilson, D., 1990, Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin, *J. Biol. Chem.* **265:**15286–15293.

Gerken, T. A., Butenhof, K. J., and Shogren, R., 1989, Effects of glycosylation on the conformation and dynamics of O-linked glycoproteins: Carbon-13 NMR studies of ovine submaxxilary mucin, *Biochemistry* **28:**5536–5543.

Gottschalk, A., 1960, Correlation between composition, structure, shape and function of a salivary mucoprotein, *Nature* **186:**949–951.

Gottschalk, A., and Thomas, M.A.W., 1961, Studies on mucoproteins. V. The significance of N-acetylneuraminic acid for the viscosity of ovine submaxillary gland glycoprotein, *Biochim. Biophys. Acta* **46:**91–98.

Green, E. D., van Halbeek, H., Boime, I., and Baenziger, J. U., 1985, Structural elucidation of the disulfated oligosaccharide from bovine lutropin, *J. Biol. Chem.* **260:**15623–15630.

Griffiths, B., Matthews, D. J., West, L., Attwood, J., Povey, S., Swallow, D. M., Gum, J. R., and Kim, Y. S., 1990, Assignment of the polymorphic intestinal mucin gene (Muc2) to chromosome-11p15, *Ann. Hum. Genet.* **54:**270–277.

Gum, J. R., Byrd, J. C., Hicks, J. W., Toribara, N. W., Lamport, D.T.A., and Kim, Y. S., 1989, Molecular cloning of human intestinal mucin cDNA's, *J. Biol. Chem.* **264:**6480–6487.

Gum, J. R., Hicks, J. W., Swallow, D. M., Lagac, R. L., Byrd, J. C., Lamport, D.T.A., Siddiki, B., and Kim, Y. S., 1990, Molecular cloning of cDNAs from a novel human intestinal mucin gene, *Biochem. Biophys. Res. Commun.* **171:**407–415.

Gum, J. R., Jr., Hicks, J. W., Toribara, N. W., Rothe, E.-M., Lagace, R. E., and Kim, Y. S., 1993, The human MUC2 intestinal mucin has cysteine-rich subdomains located both upstream and downstream of its central repetitive region, *J. Biol. Chem.* **267**:21375–21383.

Hakomori, S., 1973, Glycolipids of tumor membranes, *Adv. Cancer Res.* **18**:265–315.

Hakamori, S., 1984, Tumor-associated carbohydrate antigens, *Annu. Rev. Immunol.* **2**:103–126.

Hakomori, S., 1989, Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens, *Adv. Cancer Res.* **52**:257–331.

Hanisch, F., Pete-Katalinic, J., Egge, H., Dabrowski, U., and Uhlenbruck, G., 1990, Structures of acidic O-linked polylactosaminoglycans on human skim milk mucins, *Glycoconjugate J.* **7**:525–543.

Hanski, C., Drechsler, K., Hanisch, F.-G., Sheehan, H., Manske, M., Ogorek, D., Klussmann, E., Hanski, M.-L., Blank, M., Xing, P.-X., McKenzie, I.F.C., Devine, P. L., and Riecken, E.-O., 1993, Altered glycosylation of the MUC-1 protein core contributes to the colon carcinoma-associated increase of mucin-bound sialyl-Lewis expression, *Cancer Res.* **53**:4082–4088.

Hayes, D. F., Silberstein, D. S., Rodrique, S. W., and Kufe, D. W., 1990, DF3 antigen, a human epithelial cell mucin, inhibits adhesion of eosinophils to antibody-coated targets, *J. Immunol.* **145**:960–962.

Hession, C., Osborn, L., Goff, D., Chi-Rosso, G., Vassallo, C., Pasek, M., Pittack, C., Tizard, R., Goelz, S., McCarthy, K., Hopple, S., and Lobb, R., 1990, Endothelial leukocyte adhesion molecule 1: Direct expression cloning and functional interactions, *Proc. Natl. Acad. Sci. USA* **87**:1673–1677.

Hilgers, J., Zotter, S., and Kenemans, P., 1989, Polymorphic epithelial mucin and CA 125-bearing glycoprotein in basic and applied carcinoma research, *Cancer Rev.* **11–12**:3–10.

Hilkens, J., 1988, Biochemistry and function of mucins in malignant disease, *Cancer Rev.* **11–12**:25–54.

Hilkens, J., and Buijs, F., 1988, Biosynthesis of Mam-6, an epithelial sialomucin, *J. Biol. Chem.* **263**:4215–4222.

Hilkens, J., Ligtenberg, M.J.L., Vos, H. L., and Litvinov, S. V., 1992, Cell membrane associated mucins and adhesion-modulating property, *Trends Biochem. Sci.* **17**:359–363.

Hill, H. D., Reynolds, J. A., and Hill, R. L., 1977a, Purification, composition, molecular weight, and subunit structure of ovine submaxillary mucin, *J. Biol. Chem.* **252**:3791–3798.

Hill, H. D., Schwyzer, M., Steinman, H. M., and Hill, R. L., 1977b, Ovine submaxillary mucin—Primary structure and peptide substrates of UDP-N-acetylgalactosamine: Mucin transferase, *J. Biol. Chem.* **252**:3799–3804.

Hiraizumi, S., Takasaki, S., Shiroki, K., Kochibe, N., and Kobata, A., 1990, Transfection with fragments of the adenovirus 12 gene induces tumorigenicity-associated alteration of N-linked sugar chains in rat cells, *Arch. Biochem. Biophys.* **280**:9–19.

Hiraizumi, S., Takasaki, S., Ohuchi, N., Harada, Y., Nose, M., Shozo, M., and Kobata, A., 1992, Altered glycosylation of membrane glycoproteins associated with human mammary carcinoma, *Jpn. J. Cancer Res.* **83**:1063–1072.

Hironaka, T., Furukawa, K., Esmon, P. C., Fournel, M. A., Sawada, S., Kata, M., Minaga, T., and Kobata, A., 1992, Comparative study of the sugar chains of factor Viii purified from human plasma and from the culture media of recombinant baby hamster kidney cells, *J. Biol. Chem.* **267**:8012–8020.

Hoff, S. D., Matsushita, Y., Ota, D. M., Cleary, K. R., Yamori, T., Hakomori, S., and Irimura, T., 1989, Increased expression of sialyl-dimeric Le antigen in liver metastases of human colorectal carcinoma, *Cancer Res.* **49**:6883–6888.

Hoffmann, W., 1988, A new repetitive protein from *Xenopus laevis* skin highly homologous to pancreatic spasmolytic polypeptide, *J. Biol. Chem.* **263**:7686–7690.

Hull, S. R., Bright, A., Carraway, K. L., Abe, M., Hayes, D. F., and Kufe, D. W., 1989,

Oligosaccharide difference in the DF3 sialomucin antigen from normal human milk and the BT-20 human breast carcinoma cell line, *Cancer Commun.* **1:**261–267.

Humphries, M. J., Matsumoto, K., White, S. L., and Olden, K., 1986, Oligosaccharide modification by swainsonine treatment inhibits pulmonary colonization by B16-F10 murine melanoma cells, *Proc. Natl. Acad. Sci. USA* **83:**1752–1756.

Ingram, G. A., 1980, Substances involved in the natural resistance of fish to infection: A review, *J. Fish Biol.* **16:**23–30.

Irimura, T., Gonzalez, R., and Nicolson, G. L., 1981, Effects of tunicamycin on B16 metastatic melanoma cell surface glycoproteins and blood-borne arrest and survival properties, *Cancer Res.* **41:**3411–3418.

Irimura, T., Carlson, D. A., Price, J., Yamori, T., Giavazzi, R., Ota, D. M., and Cleary, K. R., 1988, Differential expression of a sialoglycoprotein with an approximate molecular weight of 900,000 on metastatic human colon carcinoma cells growing in culture and in tumor tissues, *Cancer Res.* **48:**2353–2360.

Irimura, T., McIsaac, A. M., Carlson, D. A., Yagita, M., Grimm, E. A., Menter, D. G., Ota, D. M., and Cleary, K. R., 1990, Soluble factor in normal tissues that stimulates high-molecular-weight sialoglycoprotein production by human colon carcinoma cells, *Cancer Res.* **50:**3331–3338.

Irimura, T., Nakamori, S., Matsushita, Y., Taniuchi, Y., Todoroki, N., Tsuji, T., Izumi, Y., Kawamura, Y., Hoff, S. D., Cleary, K. R., and Ota, D. M., 1993, Colorectal cancer metastasis determined by carbohydrate-mediated cell adhesion: Role of sialyl-Lex antigens, *Cancer Biol.* **4:**319–324.

Itzkowitz, S. H., Bloom, E. J., Kokal, W. A., Modin, G., Hakomori, S., and Kim, Y. S., 1990, A novel mucin antigen associated with prognosis in colorectal cancer patients, *Cancer* **66:**1960–1966.

Jerome, K. R., Barnd, D. L., Bendt, K. M., Boyer, C. M., Taylor-Papadimitriou, J., McKenzie, I.F.C., Bast, R. C., and Finn, O. J., 1991, Cytotoxic T-lymphocytes derived from patients with breast adenocarcinoma recognize an epitope present on the protein core of mucin molecule preferentially expressed by malignant cells, *Cancer Res.* **51:**2908–2916.

Johnson, D. C., and Spear, P. G., 1983, O-linked oligosaccharides are acquired by herpes simplex virus glycoproteins in the Golgi apparatus, *Cell* **32:**987–999.

Johnson, V. G., Schlom, J., Paterson, A. J., Bennett, J., Magnani, J. L., and Colcher, D., 1986, Analysis of a human tumor-associated glycoprotein (Tag-72) identified by monoclonal antibody B72.3, *Cancer Res.* **46:**850–857.

Johnson, W. V., and Heath, E. C., 1986, Evidence for posttranslational O-glycosylation of fetuin, *Biochemistry* **25:**5518–5525.

Jokinen, M., Andersson, L. C., and Gahmberg, C. G., 1985, Biosynthesis of the major human red blood cell sialoglycoprotein, glycophorin A, *J. Biol. Chem.* **260:**11314–11321.

Joziasse, D. H., Schiphorst, W.E.C.M., van den Eijnden, D. H., Van Kuik, J. A., Van Halbeek, H., and Vliegenthart, J.F.G., 1987, Branch specificity of bovine colostrum CMP-sialic acid: N-acetyllactosaminide $\alpha 2 \rightarrow 6$ sialyltransferase, *J. Biol. Chem.* **262:**2025–2033.

Joziasse, D. H., Shaper, J. H., van den Eijnden, D. H., Van Tunen, A. J., and Shaper, N. L., 1989, Bovine $\alpha 1 \rightarrow 3$-galactosyltransferase: Isolation and characterization of a cDNA clone, *J. Biol. Chem.* **265:**14290–14297.

Joziasse, D. H., Shaper, J. H., Jabs, E. W., and Shaper, N. L., 1991, Characterization of an $\alpha 1 \rightarrow 3$ galactosyltransferase homologue on human chromosome 12 that is organized as a processed pseudogene, *J. Biol. Chem.* **266:**6991–6998.

Kaliner, M. A., 1991, Human nasal respiratory secretions and host defense, *Am. Rev. Respir. Dis.* **144:**S52–S56.

Kawano, T., Takasaki, S., Tao, T.-W., and Kobata, A., 1993, Altered glycosylation of $\beta 1$ integrins associated with reduced adhesiveness to fibronectin and laminin, *Int. J. Cancer* **53:**91–96.

Kawashima, H., Yamamoto, K., Osawa, T., and Irimura, T., 1993, Purification and characterization of UDP-GlcNAc:Galβ1→4Glc(NAc) β1,3-N-acetyl glucosaminyltransferase (poly-N-acetyllactosamine extension enzyme) from calf serum, *J. Biol. Chem.* **268:**27118–27126.

Khatri, I. S., Forstner, G. G., and Forstner, J. F., 1993, Suggestive evidence for two different mucin genes in rat intestine, *Biochem. J.* **294:**391–399.

Kim, Y. S., Yuan, M., Itzkowitz, S. H., Sun, Q., Kaizu, T., Palekar, A., Trump, B. F., and Hakomori, S., 1986, Expression of LeY and extended LeY blood group-related antigens in human malignant, premalignant and nonmalignant colonic tissues, *Cancer Res.* **46:**5985–5992.

Kobata, A., 1989, Altered glycosylation of surface glycoproteins in tumor cells and its clinical application, *Pigment Cell Res.* **2:**304–308.

Kornfeld, S., 1982, Oligosaccharide processing during glycoprotein biosynthesis, in: *The Glycoconjugates* (M. I. Horowitz, ed.), Academic Press, New York, Vol. III, pp. 3–23.

Kotovuori, P., Tontti, E., Pigott, R., Shepherd, M., Kisco, M., Hasegawa, A., Renkonen, R., Nortamo, P., Altieri, D. C., and Gahmberg, C. G., 1993, The vascular E-selectin binds to the leukocyte integrins CD11/CD18, *Glycobiology* **3:**131–136.

Kufe, D., Inghirami, G., Abe, M., Hayes, D., Justi-Wheeler, H., and Schlom, J., 1984, Differential reactivity of a novel monoclonal antibody (DF3) with human malignant versus benign breast tumors, *Hybridoma* **3:**223–232.

Kurosaka, A., Nakajima, H., Funakoshi, I., Matsuyama, M., Nagayo, T., and Yamashina, I., 1983, Structures of the major oligosaccharides from a human rectal adenocarcinoma glycoprotein, *J. Biol. Chem.* **258:**11594–11598.

Kurosaka, A., Kitagawa, H., Fukui, S., Numata, Y., Yamashina, I., Nakada, H., Funakoshi, I., Kawasaki, T., Ogawa, T., and Iijima, H., 1988, A monoclonal antibody that recognizes a cluster of a disaccharide, Neu5Ac2→6GalNAc, in mucin-type glycoproteins, *J. Biol. Chem.* **262:**8724–8726.

Kurosawa, N., Kojima, N., Inoue, M., Hamamoto, T., and Tsuji, S., 1994, Cloning and expression of Galβ1,3 GalNAc specific GalNAcα2,6 sialyltransferase, *J. Biol. Chem.* **269:**19048–19053.

Laferte, S., and Dennis, J. W., 1989, Purification of two glycoproteins expressing β1-6 branched Asn-linked oligosaccharides from metastatic tumor cells, *Biochem. J.* **259:**569–576.

Lan, M. S., Hollingsworth, M. A., and Metzgar, R. S., 1990, Polypeptide core of a human pancreatic tumor mucin antigen, *Cancer Res.* **50:**2997–3001.

Larsen, R. D., Rajan, V. P., Ruff, M. M. Kukowsaka-Latallo, J., Cummings, R. D., and Lowe, J. B., 1989, Isolation of a cDNA encoding a murine UDP galactose:beta-D-galactosyl-1,4-N-acetyl-D-glucosaminide alpha-1,3-galactosyltransferase: Expression cloning by gene transfer, *Proc. Natl. Acad. Sci. USA* **86:**8227–8331.

Larsen, R. D., Rivera-Marrero, C. A., Ernst, L. K., Cummings, R. D., and Lowe, J. B., 1990, Frameshift and nonsense mutations in a human genomic sequence homologous to a murine Udp-Gal:Beta-D-Gal(1,4)-D-GlcNAc alpha(1,3)-galactosyltransferase cDNA, *J. Biol. Chem.* **265:** 7055–7061.

Ligtenberg, M.J.L., Vos, H. L., Gennissen, A.M.C., and Hilkens, J., 1990, Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene encoding splice variants with alternative termini, *J. Biol. Chem.* **265:**5573–5578.

Ligtenberg, M.J.L., Buijs, F., Vos, H. L., and Hilkens, J., 1992, Suppression of cellular aggregation by high levels of epsialin, *Cancer Res.* **52:**2318–2324.

Lin, L.-H., Stern, J. L., and Davidson, E. A., 1982, Clones from cultured, B16 mouse-melanoma cells resistant to wheat germ agglutinin and with altered production of mucin-type glycoproteins, *Carbohydr. Res.* **111:**257–271.

Lippincott-Schwartz, J., and Fambrough, D. M., 1987, Cycling of the integral membrane glycoprotein, LEP100, between plasma membrane and lysosomes: Kinetic and morphological analysis, *Cell* **49:**669–677.

Lis, H., and Sharon, N., 1993, Protein glycosylation. Structure and functional aspects—Review, *Eur. J. Biochem.* **218:**1–27.

Maeji, N. J., Inoue, Y., and Chujo, R., 1987, Preliminary communication: The role of the N-acetyl group in determining the conformation of 2-acetamido-2-deoxy-D-galactopyranosyl-threonine-containing peptides, *Carbohydr. Res.* **162:**4–8.

Magnani, J. L., Steplewski, Z., Koprowski, H., and Ginsburg, V., 1983, Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19-9 in the sera of patients as a mucin, *Cancer Res.* **43:**5489–5492.

Marianne, T., Perini, J. M., Lafitte, J., Houdret, N., Pruvot, F. R., Lamblin, G., Slayter, H. S., and Roussel, P., 1987, Peptides of human bronchial mucus glycoproteins, *Biochem. J.* **248:**189–195.

Matsushita, Y., Hoff, S. D., Nudelman, E. D., Ohtaka, M., Hakomori, S., Ota, D. M., Cleary, K. R., and Irimura, T., 1991, Metastatic behavior and cell surface properties of HT-29 human colon carcinoma variant cells selected for their differential expression of sialyl-dimeric LeX antigen, *Clin. Exp. Metast.* **9:**283–299.

Matsuura, H., Greene, T., and Hakomori, S., 1989, An α-N-acetyl galactosaminylation at the threonine residue of a defined peptide sequence creates the oncofetal peptide epitope in human fibronectin, *J. Biol. Chem.* **264:**10472–10476.

Medof, M. E., Lublin, D. M., Holers, V. M., Ayers, D. J., Getty, R. R., Leykam, J. K., Atkinson, J. P., and Tykocinski, M. L., 1987, Cloning and characterization of cDNAs encoding the complete sequence of decay-accelerating factor of human complement, *Proc. Natl. Acad. Sci. USA* **84:**2007–2011.

Miller, S. C., Hay, E. D., and Codington, J. F., 1977, Ultrastructural and histochemical differences in cell surface properties of strain specific and non-strain specific TA3 adenocarcinoma cells, *J. Cell Biol.* **72:**511–529.

Mizouchi, T., Yamashita, K., Fukikawa, K., Titani, K., and Kobata, A., 1980, The structures of the carbohydrate moieties of bovine blood coagulation factor X, *J. Biol. Chem.* **255:**3526–3531.

Murthy, V.L.N., and Horowitz, M. I., 1966, Alkali-reductive cleavage of ovine submaxillary mucin, *Carbohydr. Res.* **6:**266–275.

Nakada, H., Funakoshi, I., and Yamashina, I., 1975, The isolation and characterization of glycopeptides and mucopolysaccharides from plasma membranes of an ascites hepatoma, AH 130, *J. Biochem.* **78:**863–872.

Nakada, H., Numata, Y., Inoue, M., Tanaka, N., Kitagawa, H., Funakoshi, I., Fukui, S., and Yamashina, I., 1991, Elucidation of an essential structure recognized by an anti-GalNAcα-Ser (Thr) monoclonal antibody (MLS 128), *J. Biol. Chem.* **266:**12402–12405.

Nakada, H., Inoue, M., Numata, Y., Tanaka, N., Funakoshi, I., Fukui, S., Mellors, A., and Yamashina, I., 1993, Epitopic structure of Tn glycophorin A for an anti-Tn antibody (MLS 128), *Proc. Natl. Acad. Sci. USA* **90:**2495–2499.

Nakamori, S., Ota, D. M., Cleary, K. R., Shirotani, K., and Irimura, T., 1994, MUC1 mucin expression as a marker of progression and metastasis of human colorectal carcinoma, *Gastroenterology* **106:**353–361.

Nieman, H., Boschek, B., Evans, D., Rosig, M., Tamura, T., and Klenk, H., 1982, Post-translational glycosylation of coronavirus glycoprotein E1—Inhibition by monensin, *EMBO J.* **1:**1499–1504.

Okumura, T., Lombart, C., and Jamieson, G. A., 1976, Platelet glycocalicin II: Purification and characterization, *J. Biol. Chem.* **251:**5950–5955.

Parekh, R. B., Tse, A.G.D., Dwek, R. A., Williams, A. F., and Rademacher, T. W., 1987, Tissue-specific N-glycosylation, site specific oligosaccharide patterns and lentil lectin recognition of rat Thy-1, *EMBO J.* **6:**1233–1244.

Parry, G., Beck, J. C., Moss, L., Barley, J., and Ojakian, G. K., 1990, Determination of apical

membrane polarity in mammary epithelial cell cultures: The role of cell–cell, cell–substratum, and membrane–cytoskeleton interactions, *Exp. Cell Res.* **188:**302–311.

Passaniti, A., and Hart, G. W., 1988, Cell surface sialylation and tumor metastasis, *J. Biol. Chem.* **263:**7591–7603.

Peat, N., Gendler, J. J., Lalani, E., Duhig, T., and Taylor-Papadimitriou, J., 1992, Tissue-specific expression of a human polymorphic epithelial mucin (MUC1) in transgenic mice, *Cancer Res.* **52:**1954–1960.

Pierce, M., and Arango, J., 1986, Rous sarcoma virus-transformed baby hamster kidney cells express higher levels of asparagine-linked tri- and tetra-antennary glycopeptides containing [GlcNAc-β(1,6)Man-α(1,6)Man] and poly-N-acetyllactosamine sequences than baby hamster kidney cells, *J. Biol. Chem.* **261:**10772–10777.

Pierce, M., Arango, J., Tahir, S. H., and Hindsgaul, O., 1987, Activity of UDP-GlcNAc: Alpha-mannoside beta (1,6)N-acetylglucosaminyltransferase (GnT V) in cultured cells using a synthetic trisaccharide acceptor, *Biochem. Biophys. Res. Commun.* **146:**679–684.

Pigman, W., 1977, Submaxillary and sublingual glycoprotein, in: *The Glycoconjugates* (M. I. Horowitz and W. Pigman, eds.), Academic Press, New York, Vol. 1, pp. 129–135.

Piller, F., and Cartron, J.-P., 1983, UDP-GlcNAc:Galβ1→4GlcNAcβ1→3 N-acetyl glucosaminyltransferase, *J. Biol. Chem.* **258:**12293–12299.

Piller, V., Piller, F., Klier, F. G., and Fukuda, M., 1989, O-glycosylations of leukosialin in K562 cells, *Eur. J. Biochem.* **183:**123–135.

Pisano, A., Redmond, J. W., Williams, K. L., and Gooley, A. A., 1993, Glycosylation sites identified by solid phase Edman degradation: O-linked glycosylation motifs on human glycophorin A, *Glycobiology* **3:**429–435.

Porchet, N., Nguyen, V. C., Dufosse, J., Audie, J. P., Guyonnet-Duperat, V., Gross, M. S., Denis, C., Degand, P., Bernheim, A., and Aubert, J. P., 1991, Molecular cloning and chromosomal localization of a novel human tracheobronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs, *Biochem. Biophys. Res. Commun.* **175:**414–422.

Probst, J. C., Hauser, F., Joba, W., and Hoffmann, W., 1992, The polymorphic integumentary mucin B.1 from Xenopus laevis contains the short consensus repeat, *J. Biol. Chem.* **267:**6310–6316.

Rice, G. E., and Bevilacqua, M. P., 1989, An inducible endothelial cell surface glycoprotein mediates melanoma adhesion, *Science* **246:**1303–1306.

Roth, J., 1984, Cytochemical localization of terminal N-acetylgalactosamine residues in cellular compartments of intestinal goblet cells: Implications of topology of O-glycosylation, *J. Cell Biol.* **98:**399–406.

Saito, H., Nishikawa, A., Gu, J., Ihara, Y., Soejima, H., Wada, Y., Sekiya, C., Niikawa, N., and Taniguchi, N., 1994, cDNA cloning and chromosomal mapping of human N-acetylglucosaminyltransferase V, *Biochem. Biophys. Res. Commun.* **198:**318–327.

Sako, D., Chang, X.-J., Barone, K. M., Vachino, G., White, H. M., Shaw, G., Veldman, G. M., Bean, K. M., Ahren, T. J., Furle, B., Cumming, D. A., and Larsen, G. R., 1993, Expression cloning of a functional glycoprotein ligand for P-selectin, *Cell* **75:**1179–1186.

Sanford, B. H., 1967, An alteration in tumor histocompatibility induced by neuraminidase, *Transplantation* **5:**1273–1279.

Santer, U. V., Gilbert, F., and Glick, M. C., 1984, Change in glycosylation of membrane glycoproteins after transfection of NIH 3T3 with human tumor DNA, *Cancer Res.* **44:**3730–3735.

Sasaki, K., Watanabe, E., Kawashima, K., Sekine, S., Dohi, T., Oshima, M., Hanai, N., Nishi, T., and Hasegawa, M., 1993, Expression cloning of a novel Galβ(1→3/1→4)GlcNAc α2,3-sialyltransferase using lectin resistance selection, *J. Biol. Chem.* **268:**22782–22787.

Sato, T., Furukawa, K., Autero, M., Gahmberg, C. G., and Kobata, A., 1993a, Structural study of the sugar chains of human leukocyte common antigen CD45, *Biochemistry* **32:**12694–12703.

Sato, T., Furukawa, K., Greenwalt, D. E., and Kobata, A., 1993b, Most bovine milk fat globule

membrane glycoproteins contain asparagine-linked sugar chains with GalNAcβ1→4GlcNAc groups, *J. Biol. Chem.* **114:**890–900.

Sawada, R., Lowe, J. B., and Fukuda, M., 1993, E-selectin-dependent adhesion efficiency of colonic carcinoma cells is increased by genetic manipulation of their cell surface lysosomal membrane glycoprotein-1 expression levels, *J. Biol. Chem.* **268:**12675–12681.

Sawada, R., Tsuboi, S., and Fukuda, M., 1994, Differential E-selectin-dependent adhesion efficiency in sublines of a human colon cancer exhibiting distinct metastatic potentials, *J. Biol. Chem.* **269:**1425–1431.

Sawada, T., Ho, J.J.L., Sagabe, T., Yoon, W.-H., Chung, Y.-S., Sowa, M., and Kim, Y. S., 1993, Biphasic effect of cell surface sialic acids on pancreatic cancer cell adhesiveness, *Biochem. Biophys. Res. Commun.* **195:**1096–1103.

Schachter, H., 1986, Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides, *Biochem. Cell Biol.* **64:**163–181.

Schachter, H., and Brockhausen, I., 1992, The biosynthesis of serine (threonine)-N-acetylgalactosamine-linked carbohydrate moieties, in: *Glycoconjugates—Composition, Structure and Function* (H. J. Allen and E. C. Kisailus, eds.), Dekker, New York, pp. 263–404.

Schweizer, A., Clausen, H., Van Meer, G., and Hauri, H.-P., 1994, Localization of O-glycan initiation, sphingomyelin synthesis, and glucosylceramide synthesis in Vero cells with respect to the endoplasmic reticulum–Golgi intermediate compartment, *J. Biol. Chem.* **269:**4035–4041.

Segawa, K., and Yamaguchi, N., 1986, Characterization of the chimeric SV40 large T antigen which has a membrane attachment sequence of polyoma virus middle T antigen, *Virology* **155:**334–344.

Sgroi, D., Varki, A., Braesch-Andersen, S., and Stamenkovic, I., 1993, CD22, a B cell-specific immunoglobulin superfamily member, is a sialic acid-binding lectin, *J. Biol. Chem.* **268:**7011–7018.

Sheng, Z., Wu, K., Carraway, K. L., and Fregien, N., 1992, Molecular cloning of the transmembrane component of the 13762 mammary adenocarcinoma sialomucin complex, *J. Biol. Chem.* **267:**16341–16446.

Sherblom, A. P., and Carraway, K. L., 1980, Sulfate incorporation into the major sialoglycoprotein of the MAT-B1 subline of the 13762 rat ascites mammary adenocarcinoma, *Biochemistry* **19:**1213–1219.

Sherblom, A. P., Buck, R. L., and Carraway, K. L., 1980a, Purification of the major sialoglycoproteins of 13762 MAT-B1 and MAT-C1 rat ascites mammary adenocarcinoma cells by density gradient centrifugation in cesium chloride and guanidine hydrochloride, *J. Biol. Chem.* **255:**783–790.

Sherblom, A. P., Huggins, J. W., Chesnut, R. W., Buck, R. L., Ownby, C. L., Dermer, G. B., and Carraway, K. L., 1980b, Cell surface properties of ascites sublines of the 13762 rat mammary adenocarcinoma, *Exp. Cell Res.* **126:**417–426.

Shimizu, M., and Yamauchi, K., 1982, Isolation and characterization of mucin-like glycoprotein in human milk fat globule membrane, *J. Biol. Chem.* **91:**515–524.

Shimizu, M., Tanimoto, H., Azuma, N., and Yamauchi, K., 1990, Growth inhibition of Balb/C 3T3 cells by a high-molecular-weight mucin-like glycoprotein of human milk fat globule membrane, *Biochem. Int.* **20:**147–154.

Shimizu, Y., and Shaw, S., 1993, Mucins in the mainstream, *Nature* **366:**630–631.

Shirotani, K., Taylor-Papadimitriou, J., Gendler, S. J., and Irimura, T., 1994, Transcriptional regulation of MUC1 gene in colon carcinoma cells by a soluble factor: Identification of a regulatory element, *J. Biol. Chem.* **269:**15030–15035.

Shogren, R., Gerken, T. A., and Jentoft, N., 1989, Role of glycosylation on the conformation and chain dimensions of O-linked glycoproteins: Light-scattering studies of ovine submaxillary mucin, *Biochemistry* **28:**5525–5536.

Shoreibah, M., Perng, G.-S., Adler, B., Weinstein, J., Basu, R., Cupples, R., Wen, D., Browne, J. K., Buckhaults, P., Fregien, N., and Pierce, M., 1993, Isolation, characterization, and expression of a cDNA encoding N-acetyl glucosaminyltransferase V, *J. Biol. Chem.* **268:** 15381–15385.

Smith, P. L., and Baenziger, J. U., 1992, Molecular basis of recognition by the glycoprotein hormone-specific N-acetylgalactosamine-transferase, *Proc. Natl. Acad. Sci. USA* **89:**329–333.

Spicer, A. P., Parry, G., Patton, S., and Gendler, S. J., 1991, Molecular cloning and analysis of the mouse homologue of the tumor-associated mucin, MUC1, reveals conservation of potential O-glycosylation sites, transmembrane, and cytoplasmic domains and a loss of minisatellite-like polymorphism, *J. Biol. Chem.* **266:**15099–15109.

Spielman, J., Hull, S. R., Sheng, Z., Kanterman, R., Bright, A., and Carraway, K. L., 1988, Biosynthesis of a tumor cell surface sialomucin, *J. Biol. Chem.* **263:**9621–9629.

Springer, G. F., 1984, T and Tn, general carcinoma autoantigens, *Science* **224:**1198–1206.

Spooncer, E., Fukuda, M., Klock, J. C., Oates, J. E., and Dell, A., 1984, Isolation and characterization of polyfucosylated lactosaminoglycan from human granulocytes, *J. Biol. Chem.* **259:**4792–4801.

Stanbridge, E. J., Der, C. J., Doerson, C., Nishimi, R. Y., Wilkinson, J. E., Peehl, D. M., and Weissman, B. E., 1982, Human cell hybrids: Analysis of transformation and tumorigenicity, *Science* **215:**252–259.

Steck, P. A., and Nicolson, G. L., 1983, Cell surface glycoproteins of 13762Nf mammary adenocarcinoma clones of differing metastatic potentials, *Exp. Cell Res.* **147:**255–267.

Strous, G. J., and Dekker, J., 1992, Mucin-type glycoproteins, *Crit. Rev. Biochem. Mol. Biol.* **27:**57–92.

Strous, G.J.A., 1979, Initial glycosylation of proteins with acetylgalactosaminyl serine linkages, *Proc. Natl. Acad. Sci. USA.***76:**2694–2698.

Swallow, D. M., Griffiths, B., Bramwell, M., Wiseman, G., and Burchell, J., 1986, Detection of the urinary 'pum' polymorphism by the tumor-binding monoclonal antibodies Ca1, Ca2, Ca3, HMFG1, and HMFG2, *Dis. Markers* **4:**247–254.

Swallow, D. M., Gendler, S., Griffiths, B., Kearney, A., Povey, S., Sheer, D., Palmer, R. W., and Taylor-Papadimitriou, J., and Bramwell, M., 1987a, The hypervariable gene locus PUM, which codes for the tumor associated epithelial mucins, is located on chromosome 1, within the region 1q21-24, *Ann. Hum. Genet.* **51:**289–294.

Swallow, D. M., Gendler, S., Griffiths, B., Corney, G., Taylor-Papadimitriou, J., and Bramwell, M. E., 1987b, The human tumor-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM, *Nature* **328:**82–84.

Symond, D. L., and Vickery, A. L., 1976, Mucinous carcinoma of the colon and rectum, *Cancer* **37:**1891–1900.

Takeuchi, M., Takasaki, S., Miyazaki, H., Kato, T., Hoshi, S., Kochibe, N., and Kobata, A., 1988, Comparative study of the asparagine-linked sugar chains of human erythropoietins purified from urine and the culture medium of recombinant Chinese hamster ovary cells, *J. Biol. Chem.* **263:**3657–3663.

Tao, T.-W., and Johnson, L. K., 1982, Altered adhesiveness of tumor cell surface variants with reduced metastasizing capacity. Reduced adhesiveness to vascular wall components in culture, *Int. J. Cancer* **30:**763–766.

Thomas, M. L., 1989, The leukocyte common antigen family, *Annu. Rev. Immunol.* **7:**339–369.

Timpte, C. S., Eckhardt, A. E., Abernethy, J. L., and Hill, R. L., 1988, Porcine submaxillary gland apomucin contains tandemly repeated, identical sequences of 81 residues, *J. Biol. Chem.* **263:**1081–1088.

Tomita, M., and Marchesi, V. T., 1975, Amino-acid sequence and oligosaccharide attachment sites of human erythrocyte glycophorin, *Proc. Natl. Acad. Sci. USA* **72:**2964–2968.

Tomiya, N., Awaya, J., Kurono, M., Hanzawa, H., Shimada, I., Arata, Y., Yoshida, T., and Takahashi, N., 1993, Structural elucidation of a variety of GalNAc-containing N-linked oligosaccharides from human urinary kallidinogenase, *J. Biol. Chem.* **268:**113–126.

Toribara, N. W., Gum, J. R., Culhane, P. J., Lagace, R. E., Hicks, J. W., Petersen, G. M., and Kim, Y. S., 1991, MUC-2 human small intestinal mucin gene structure, *J. Clin. Invest.* **88:**1005–1013.

Toribara, N. W., Roberton, A. M., Ho, S. B., Kuo, W.-L., Gum, E., Hicks, J. W., Gum, J. R., Jr., Byrd, J. C., Siddiki, B., and Kim, Y. S., 1993, Human gastric mucin, *J. Biol. Chem.* **268:**5879–5885.

Tsuda, T., Gallup, M., Jany, B., Gum, J., Kim, Y. S., and Basbaum, C., 1993, Characterization of a rat airway cDNA encoding a mucin-like protein, *Biochem. Biophys. Res. Commun.* **195:**363–373.

van den Eijnden, D. H., Evans, N. A., Codington, J. F., Reinhold, V., Silber, C., and Jeanloz, R. W., 1979, Chemical structure of epiglycanin, the major glycoprotein of the TA3-Ha ascites cell, *J. Biol. Chem.* **254:**12153–12159.

van den Eijnden, D. H., Koenderman, A.H.L., and Schiphorst, W.E.C.M., 1988, Biosynthesis of blood group i-active polylactosaminoglycans, *J. Biol. Chem.* **263:**12461–12471.

van de Wiel-van Kemenade, E., Ligtenberg, M.J.L., de Boer, A. J., Buijs, F., Vos, H. L., Melief, C.J.M., Hilkens, J., and Figdor, C. G., 1993, Episialin (MUC1) inhibits cytotoxic lymphocyte–target cell interaction, *J. Immunol.* **151:**767–776.

Van Nieuw Amerongen, A., Oderkerk, C. H., Roukema, P. A., Wolf, J. H., Lisman, J.J.W., and Vliegnthart, J.F.G., 1987, Primary structure of O- and N-glycosidic carbohydrate chains derived from murine submandibular mucin. *Carbohydr. Res.* **164:**43–51.

Varki, A., 1993, Biological roles of oligosaccharides: All of the theories are correct, *Glycobiology* **3:**97–130.

Verma, M., and Davidson, E. A., 1993, Molecular cloning sequencing of a canine tracheobronchial mucin cDNA containing a cysteine-rich domain, *Proc. Natl. Acad. Sci. USA* **90:**7144–7148.

Wang, Y., Agarwal, N., Eckhardt, A. E., Stevens, R. D., and Hill, R. L., 1993, The acceptor substrate specificity of porcine submaxillary UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is dependent on the amino acid sequences adjacent to serine and threonine residues, *J. Biol. Chem.* **268:**22979–22983.

Watanabe, K., Hakomori, S., Childs, R. A., and Feizi, T., 1979, Characterization of a blood group I-active ganglioside, *J. Biol. Chem.* **254:**3221–3228.

Watkins, S. C., Slayter, H. S., and Codington, J. F., 1991, Intracellular pathway of a mucin-type membrane glycoprotein in mouse mammary tumor cells, *Carbohydr. Res.* **213:**185–200.

Weinstein, J., de Souza-e-Silva, U., and Paulson, J. C., 1982, Sialylation of glycoprotein oligosaccharides N-linked to asparagine, *J. Biol. Chem.* **257:**13845–13853.

Weinstein, J., Lee, E. U., McEntee, K., Lai, P.-H., and Paulson, J. C., 1987, Primary structure of β-galactoside 2,6-sialyltransferase, *J. Biol. Chem.* **262:**17735–17743.

Wen, D. X., Livingston, B. D., Medzihradszky, K. F., Kelm, S., Burlingame, A. L., and Paulson, J. C., 1992, Primary structure of Galβ1,2(4)GlcNAc α2,3-sialyltransferase determined by mass spectrometry sequence analysis and molecular cloning, *J. Biol. Chem.* **267:**21011–21019.

Wieruszeski, J. M., Michalski, J. C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., Van Halbeek, H., Mutsaers, J.H.G.M., and Vliegenthart, J.F.G., 1987, Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the Chinese switlet, *J. Biol. Chem.* **262:**6650–6658.

Williams, C. J., Wreschner, D. H., Tanaka, A., Tsarfaty, I., Keydar, I., and Dion, A. S., 1990, Multiple protein forms of the human breast tumor-associated epithelial membrane antigen (EMA) are generated by alternative splicing and induced by hormonal stimulation, *Biochem. Biophys. Res. Commun.* **170:**1331–1338.

Wilson, I.B.H., Gravel, Y., and Heijne, G. V., 1991, Amino acid distributions around O-linked glycosylation sites, *Biochemistry* **275:**529–534.

Woodward, H. D., Ringler, N. J., Selvakumar, R., Simet, I. M., Bhavanandan, V. P., and Davidson, E. A., 1987, Deglycosylation studies on tracheal mucin glycoproteins, *Biochemistry* **26:**5315–5322.

Xu, G., Huan, L., Khatri, J. A., Wang, D., Bennick, A., Fahim, R.E.F., Forstner, G. G., and Forstner, J., 1992, cDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide, *J. Biol. Chem.* **267:**5401–5407.

Yagel, S., Feinmesser, R., Waghorne, C., Lala, P. K., Breitman, M. L., and Dennis, J. W., 1989, Evidence that GlcNAcβ1→6 branched Asn-linked oligosaccharides on metastatic tumor cells facilitate invasion of basement membranes, *Int. J. Cancer* **44:**685–690.

Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S., 1990, Molecular genetic basis of the histo-blood group ABO system, *Nature* **345:**229–233.

Yamamura, K., Takasaki, S., Ichihashi, M., Mishima, Y., and Kobata, A., 1991, Increase of sialylated tetraantennary sugar chains in parallel to the higher lung-colonizing abilities of mouse melanoma clones, *J. Invest. Dermatol.* **97:**735:741.

Yamashita, K., Hitoi, A., Taniguchi, N., Yokosawa, N., Tsukada, Y., and Kobata, A., 1983, Comparative study of the sugar chains of δ-glutamyl transpeptidases purified from rat liver and rat AH-66 hepatoma cells, *Cancer Res.* **43:**5059–5063.

Yamashita, K., Ohkura, T., Tachibana, Y., Takasaki, S., and Kobata, A., 1984, Comparative study of the oligosaccharides released from baby hamster kidney cells and their polyoma transformant by hydrazinolysis, *J. Biol. Chem.* **259:**10834–10840.

Yamashita, K., Tachibana, Y., Ohkura, T., and Kobata, A., 1985, Enzymatic basis for the structural changes of asparagine-linked sugar chains of membrane glycoproteins of baby hamster kidney cells induced by polyoma transformation, *J. Biol. Chem.* **260:**3963–3969.

Yazawa, S., Abbas, S. A., Madiyalakan, R., Barlow, J. J., and Matta, K. L., 1986, N-acetylglucosaminyltransferases related to the synthesis of mucin-type glycoproteins in human ovarian tissue, *Carbohydr. Res.* **149:**241–242.

Yogeeswaran, G., and Tao, T., 1980, Cell surface sialic acid expression of lectin-resistant variant clones of B16 melanoma with altered metastasizing potential, *Biochem. Biophys. Res. Commun.* **95:**1452–1460.

Yonezawa, S., Byrd, J. C., Dahiya, R., Ho, J.J.L., Gum, J. R., Griffiths, B., Swallow, D. M., and Kim, Y. S., 1991, Differential mucin gene expression in human pancreatic and colon cancer cells, *Biochem. J.* **276:**599–605.

Yousefi, S., Higgins, E., Daoling, Z., Pollex-Kruger, A., Hindsgaul, O., and Dennis, J. W., 1991, Increased UDP-GlcNAc:Galβ1→3GalNAc-R(GlcNAc to GalNAc) β-1,6-N-acetylglucosaminyltransferase activity in metastatic murine tumor cell lines, *J. Biol. Chem.* **266:**1772–1782.

Zaretsky, J. Z., Weiss, M., Tsarfaty, I., Hareuveni, M., Wreschner, D. H., and Keydar, I., 1990, Expression of genes coding for Ps2, C-Erbb2, estrogen receptor and the H23 breast tumor-associated antigen. A comparative analysis in breast cancer, *FEBS Lett.* **265:**46–50.

Zotter, S., Hageman, P. C., Lossnitzer, A., Mooi, W. J., and Hilgers, J., 1988, Tissue and tumor distribution of human polymorphic epithelial mucin, *Cancer Rev.* **11–12:**55–101.