ELSEVIER

Carbohydrate Polymers 41 (2000) 293–309

Carbohydrate Polymers

www.elsevier.com/locate/carbpol

Involvement of sugars in protein–protein interactions

P.K. Qasba*

Structural Glycobiology Section, Laboratory of Experimental and Computational Biology, National Cancer Institute, P.O. Box B, Building 469, Room 221, NCI-FCRDC, Frederick, MD 21702-1201, USA

Accepted 20 September 1999

Abstract

In a cell there is an intricate machinery involved in the biosynthesis of complex oligosaccharides of glycoproteins, glycolipids, and proteoglycans. These oligosaccharides play important roles in several biological processes, including the folding and transport of glycoproteins across cellular compartments. Defective glycan synthesis has been shown to have serious pathological consequences and result in several human diseases. The oligosaccharide moieties bind to cellular proteins with high specificity and modulate the homo- and heterodimerization of glycoproteins. Owing to the conformational flexibility of oligosaccharides, the torsional angles of a disaccharide unit, particularly around the 1-6-linkage, adjust in such a way that the side groups of the oligosaccharides orient themselves in a manner that promotes favorable interactions with the binding residues of the protein. Branched oligosaccharides cross-link proteins and generate infinite networks of protein–carbohydrate complexes, resulting in the modulation of various cell responses. Glycosaminoglycans, the oligosaccharide moieties of proteoglycans, bind growth factors with a high degree of specificity and induce interactions with growth factor receptors, thereby regulating their biological activity. Using an experimental model system to study the sugar mediated protein–protein interactions, we are investigating the sugar ligand-dependent interactions between α-Lactalbumin (α-LA) and β-1,4-galactosyltransferase (β4Gal-T). Also, using a molecular modeling method, we are investigating the heparin dependent dimerization of fibroblast growth factor and growth factor receptor. Our results, together with studies from various other laboratories, support the hypothesis that in the oligosaccharide induced recognition process, sugars accelerate the assembly of the complex by positioning and orienting the molecules in an optimal fashion that brings about specific protein–protein or protein–carbohydrate interactions.

Keywords: Glycosaminoglycans; α-Lactalbumin; GalNAc residue

1. Overview

Complex carbohydrate moieties of glycoproteins and glycolipids at the cell surface are involved in the cell–cell contact, as well as cellular, bacterial and viral adhesions. Chemical structures of many of these carbohydrate moieties change during cellular development and differentiation. Certain cell-stage specific carbohydrate sequences play a key role in intercellular recognition and serve as cell surface markers. Although there is no single unifying function assigned to the oligosaccharide moiety of glycoconjugates, it is now well recognized that they aid in glycoprotein solubility and protect proteins from protease degradation. They are also the recognition molecules for lectins and, together with chaperones play an important role in the in vivo folding of glycoproteins. In addition, glycosaminoglycans, the carbohydrate moieties of proteoglycans, take part in a wide range of biological functions. For example, they are involved, together with collagen, in the formation of an extracellular matrix to which growth factors bind with a high degree of specificity and thereby regulate growth factor activity.

2. General structure of glycoconjugates

Most oligosaccharide moieties of glycoproteins are either *O*-linked or *N*-linked to proteins (Fig. 1). Generally *O*-glycosidic linkage occur between the GalNAc residue of the oligosaccharide and the side chain hydroxyl group of a serine or threonine residue of the protein. In proteoglycans, *O*-glycosidic linkage occurs between the Ser residue of the polypeptide chain and a xylose residue of the oligosaccharide. Recently many cytoplasmic and nuclear-pore proteins and transcription factors have been shown to have a GlcNAc residue linked to Ser/Thr residues of the protein (Haltiwanger et al., 1992). Fucose has also been found to be linked to Ser/Thr in the consensus sequence Gly–Gly–Ser/Thr in epidermal growth factor and several blood coagulation

* Tel.: +1-301846-1933; fax: +1-301846-7149.
E-mail address: qasba@helix.nih.gov (P.K. Qasba)

0144-8617/00/$ - see front matter
PII: S0144-8617(99)00148-4

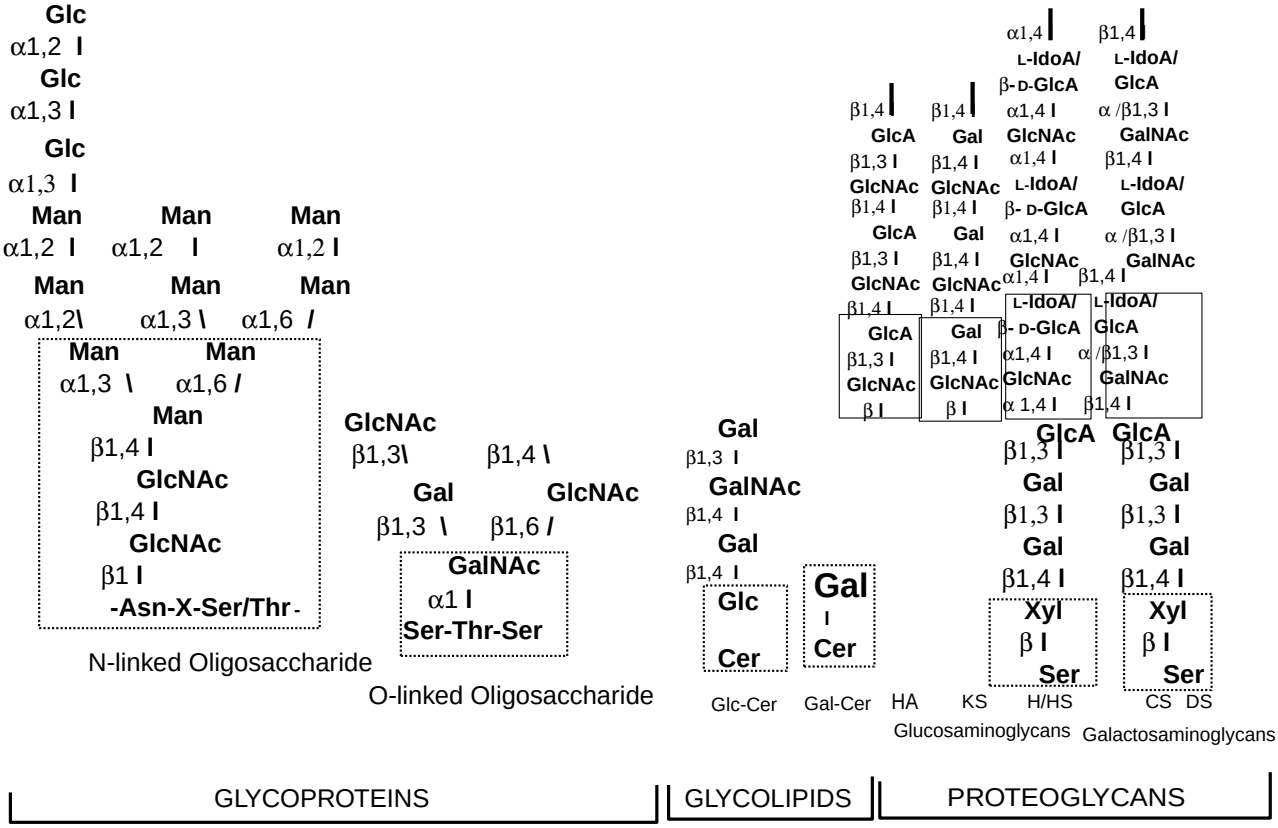

Fig. 1. Oligosaccharide Structures of Glycoconjugates. *N*-linked oligosaccharides: the side chain amide group of an asparagine residue in the consensus sequence Asn-X-Ser/Thr of a protein is linked to the GlcNAc residue of a core oligosaccharide structure, Glc3Man9GlcNac2. The core structure is processed in the Golgi apparatus where it gives rise to high mannose, complex or hybrid type structures (Fig. 2). *O*-linked oligosaccharides: Ser or Thr residue of a protein is linked to GalNAc residue of an oligosaccharide. *O*-linked core structures differ in the sequence attached to the GalNAc residue. Glycolipids: The hydroxyl group of a ceramide moiety is either connected to glucose as in glucocerebrosides or to galactose as in galactocerebrosides. Proteoglycans: Ser residue of a protein molecule is connected to a xylose residue of an oligosaccharide, which generally consists of negatively charged, repeating disaccharide units. The residues in the boxes are found at the respective linkages in various glycoproteins, glycolipids and proteoglycans, respectively.

and fibrinolytic proteins (Harris & Spellman, 1993). Although recently many other novel forms of protein glycosylation have been observed (Vliegenthart & Casset, 1998), it is the Asn-linked glycosylation that has been extensively studied in terms of structure and biosynthesis (see below).

3. Golgi-resident glycosyltransferase super-family

In the Golgi apparatus of a cell, intricate machinery is involved in the biosynthesis of complex carbohydrates on glycoproteins and glycolipids. From a small number of saccharide units, diverse oligosaccharide structures are generated through isomeric linkages between sugars. Each linkage is catalyzed by a different enzyme, and these are collectively called glycosyltransferases. These enzymes reside in the Golgi apparatus and transfer a single sugar residue from a sugar nucleotide donor to a specific oligosaccharide acceptor. Glycosyltransferases are grouped into families based on the type of sugar they transfer, e.g. *N*-acetylglucosaminyltransferase, galactosyltransferase, fucosyltransferase and sialyltransferase, transfer *N*-acetylglucosamine, galactose, fucose and sialic acid, respectively, to the acceptor substrates. Within the glycosyltransferase families, sub-families of transferases generate specific linkages, e.g. β-1,4-galactosyltransferase, β-1,3-galactosyltransferase and α-1-3-galactosyltransferase, each transfer galactose in β-1,4-, β-1,3- and α-1-3- linkage, respectively, to the acceptor molecules. After the initial cloning of milk β-1,4-galactosyltransferase from my laboratory (Narimatsu, Sinha, Brew, Okayama & Qasba, 1986) and from Shaper's laboratory (Shaper et al., 1986), many Golgi-glycosyltransferases have been cloned, and expressed, and the topologies of these enzymes predicted (reviewed in Brockhausen & Schachter, 1997). Recently, many members within the sub-families of these enzymes have been cloned and identified. The members within the sub-family show high sequence homologies with each other. For example, in the human species, six different sequence related β-1,4-galactosyltransferases have been identified, each member showing preferred acceptor specificities and tissue specific expression (Almeida et al., 1997; Lo, Shaper, Pevsner & Shaper, 1998).

Some glycosyltransferases have also been shown to be present at the cell surface interacting with the cell matrix.

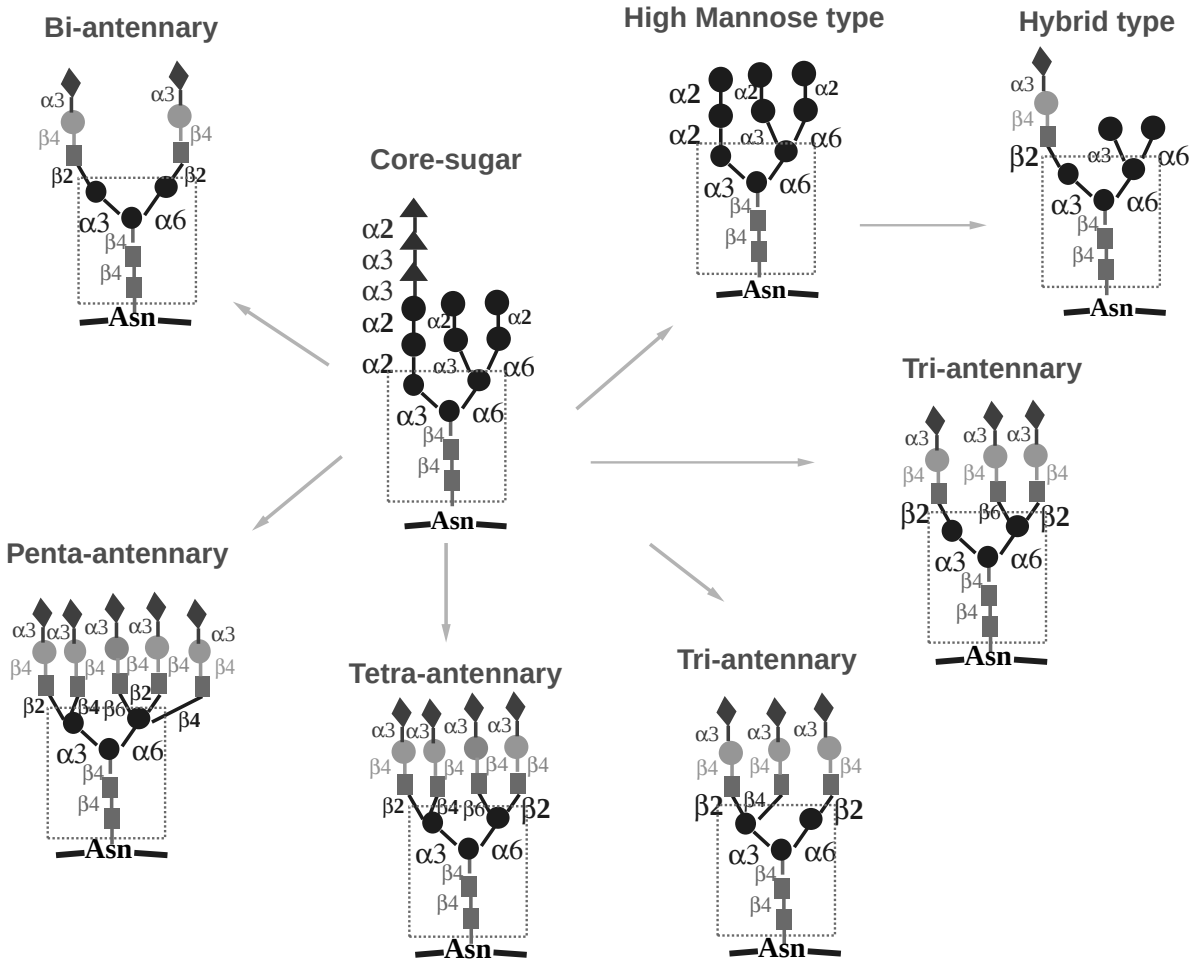

Fig. 2. Branching patterns in *N*-Glycans. The core oligosaccharide structure, Glc3Man9GlcNac2, that is attached to protein in the ER gets processed to a variety of structures in the Golgi apparatus. Common to all these structures is a core pentasaccharide, Man3GlcNAc2, shown in the dashed box. These structures may be of high mannose type, hybrid type or complex type. The complex type structure may be either of bi-, tri-, tetra-, or penta-antennary type. Linkage types and the residues are shown. Glucose (▲); Sialic acid (◆); Galactose (gray circle); *N*-Acetylglucosamine (■); Mannose (black circle).

They have been implicated as signaling molecules for pattern formation during development (Yuan, Schultz, Mlodzik & Bork, 1997). Further, several reports have indicated alterations in glycosyltransferase activities during disease conditions (see below). Hence, the knowledge of the three-dimensional structure of glycosyltransferases and their complexes with oligosaccharides is highly desirable for understanding the biosynthesis and functions of glycoproteins. The three-dimensional structure of none of the Golgi-resident glycosyltransferases is known to date.

4. Asn-linked oligosaccharides and their biosynthesis—overview

The oligosaccharide moiety in *N*-linked glycoproteins is connected to the side chain amide group of an asparagine residue in the sequence Asn-X-Ser/Thr of a protein. All the Asn-linked oligosaccharides have a common pentasaccharide core: Man-α1 → 6(Man-α1 → 3)-Man-β1 → 4-GlcNAc-β1 → 4-GlcNAc-β1 → Asn (Fig. 2). The terminal α1 → 6- and α1 → 3-linked mannose residues may carry additional residues such as mannose, *N*-acetylglucosamine, fucose, galactose, *N*-acetylgalactosamine, and sialic acid. Based on the nature of the residues present on the pentasaccharide core, Asn-linked oligosaccharides are further classified into three types: high mannose (only mannose residues), hybrid type (mannose residues on the α1 → 6-mannose and other residues on the α1 → 3-mannose) and complex type (residues other than mannose). Further, the core pentasaccharide is often either fucosylated (α1 → 3/6-linked to GlcNAc) or xylosylated (β1 → 2-linked to middle mannose), and the hybrid and complex types are sometimes bisected (β1 → 4-GlcNAc on middle mannose). Saccharides of some glycoproteins are also either sulfated or phosphorylated. The presence or absence of and the type of Asn-linked oligosaccharides have been shown to play

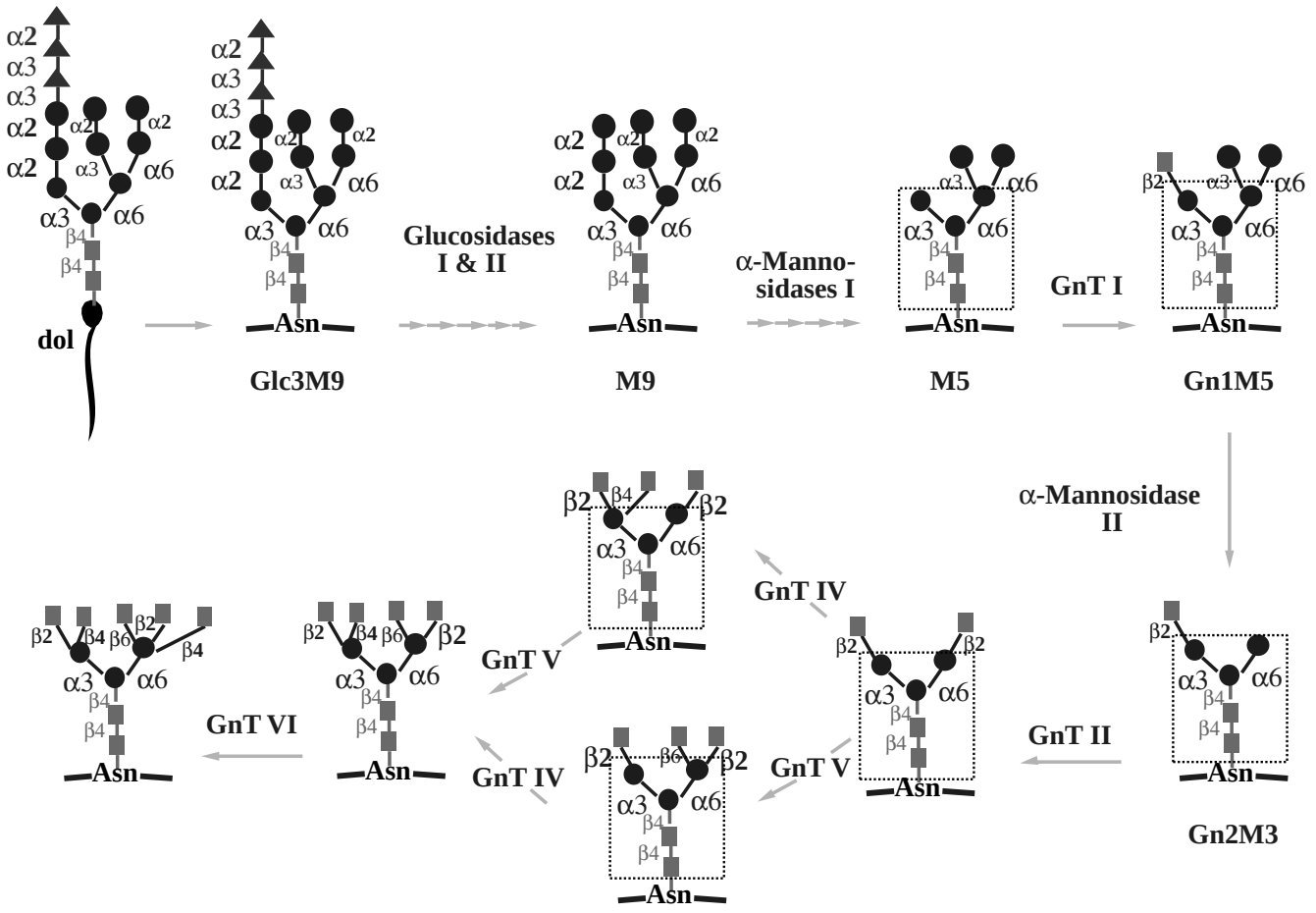

Fig. 3. Processing of Peptide-bound *N*-linked oligosaccharide chain. In the rough endoplasmic reticulum (ER) the Glc3Man9GlcNAc2 core structure from the dolichol pyrophosphate donor (dol) is transferred en bloc to an Asn residue in the sequence Asn-X-Ser/Thr of the growing polypeptide chain. The Glc3M9 structure is processed by glucosidases I and II to Man9 (M9) structure. Type I α-mannosidases remove α1-2-linked mannoses to give Man5 (M5) structure to which GlcNAc-transferase I (GnT I) transfers GlcNAc to mannose on the 1-3-arm of the core pentasaccharide (shown in the dotted box). Solid arrows indicate a single step of processing whereas dotted arrows indicate multiple steps of processing. Linkage types and the residues are shown. Glucose (▲); *N*-Acetylglucosamine (■); Mannose (●).

important roles in the biological function of many glycoproteins (Collier, Carpentier, Beitz, Carol, Taylor & Gorden, 1993; Fast, Jamieson & McCaffrey, 1993; Zhang & Dahms, 1993).

In the first step of protein *N*-glycosylation, a Glc3Man9GlcNAc2 precursor is transferred en bloc from the dolichol pyrophosphate donor to the Asn-residue of the nascent polypeptide chain (Fig. 3) (Kornfeld & Kornfeld, 1985). This co-translational event is followed by the processing of the precursor oligosaccharide by glycosidases and glycosyltransferases, leading to various types of Asn-linked oligosaccharides. The three terminal glucose residues of the Glc3Man9GlcNAc2 (Fig. 3, Glc3M9) precursor are removed by α-glucosidases forming Man9GlcNAc2 (M9), which is either left as such or is processed to different isomers of Man8GlcNAc2 by specific Man9-mannosidases (Weng & Spiro, 1993). One isomer of Man8GlcNAc2 can also be generated directly without the prior formation of Man9GlcNAc2 (Fig. 3, M9) by an endo-mannosidase which converts Glc[α1 → 3]Man9GlcNAc2 to Man8GlcNAc2 (Lubas & Spiro, 1987). Man8GlcNAc2 is further processed by α1 → 2-linkage specific mannosidase(s) I to different levels forming other high-mannose type oligosaccharide intermediates which finally lead to Man5GlcNAc2 (Fig. 3, M5). Based on several experimental observations, it has been suggested that the processing of Man9GlcNAc2 (M9) to Man5GlcNAc2 (M5) by α-mannosidase(s) I proceeds in well defined, sequential pathways (Snider, 1984), and that the accessibility and conformation of the oligosaccharides play important roles in determining these processing steps (Balaji, Qasba & Rao, 1994; Qasba & Kumar, 1997).

The GlcNAc transferase I transfers GlcNAc in β1 → 2 linkage to the core α1 → 3-mannose of Man5GlcNAc2 (Fig. 3, M5) to generate GlcNAc1Man5GlcNAc2 (Fig. 3, Gn1M5). Such an addition of GlcNAc is a necessary step for the biosynthesis of hybrid and complex type oligosaccharides. From GlcNAc1Man5GlcNAc2 (Gn1M5), either

various hybrid oligosaccharides are generated by the concerted action of several glycosyltransferases or the complex type oligosaccharide biosynthesis is initiated with the action of α-mannosidase(s) II. The α-mannosidase II removes the terminal $\alpha1 \rightarrow 3$- and $\alpha1 \rightarrow 6$-linked mannose residues on the $\alpha1 \rightarrow 6$-arm of GlcNAc1-Man5GlcNAc2 (Gn1M5), forming GlcNAc1Man3GlcNAc2 (Gn1M3) (Fig. 3). This is a committed step in the biosynthesis of complex type oligosaccharides, and this intermediate serves as the initial substrate for other glycosyltransferases such as GlcNAc-, Gal-, GalNAc-, Fuc- and Sialyl-transferases, leading to the formation of diverse complex type oligosaccharides.

Even though the assembly and processing of Asn-linked oligosaccharides follow essentially the same pathways in many organisms, the factors responsible for the nature and type of the oligosaccharide present on the glycoproteins are still not very clear. The nature of the oligosaccharide at a glycosylation site for some glycoproteins is conserved in different species (Williams et al., 1993; Zamze, Wooten, Ashford, Ferguson, Dwek & Rademacher, 1990), whereas for other glycoproteins it differs from one species to another (Dahms & Hart, 1986; Green & Baenziger, 1988). In addition to species dependency, the exact processing and hence the nature of the oligosaccharide has also been shown to depend on the tissue itself, the developmental stage, the amino acid sequence of the entire protein (i.e. the three-dimensional structure of the protein), and the specificities and levels of processing enzymes (Bischoff, Liscum & Kornfeld, 1986; Cohen & Ballou, 1980; Lubas & Spiro, 1988; Moss, Prakobphol, Wiedmann, Fisher & Damsky, 1994). The availability of the repertoire of glycosyltransferases in the cell, which in turn depends on the developmental stage of the tissue, determines the structure of the oligosaccharide moiety of a glycoprotein. Nonetheless, in a cell, the same glycosylation machinery is available to all proteins that enter the secretory pathway, yet most glycoproteins emerge with characteristic glycosylation patterns and heterogeneous populations of glycans at each glycosylation site. Individual glycoproteins in the cell, by virtue of their unique structures, can selectively control their own glycosylation by influencing where the oligosaccharide moiety of the protein and the structural motifs in the protein are recognized together by the glycosyltransferases (Manzella, Hooper & Baenziger, 1996). Spatial and steric considerations of the oligosaccharide ligands are important in determining the modes of binding and for glycosyltransferase action. Hence, a clear understanding of these sequential pathways requires detailed information about the complexes of these oligosaccharide intermediates with glycosidases and glycosyltransferases. In the absence of three-dimensional structure information of these enzymes, the conformations of the oligosaccharide intermediates alone have been studied in my laboratory, and the information obtained from these studies has been correlated with known experimental data.

5. Oligosaccharide conformations and their interactions with proteins

The oligosaccharide moiety in the carbohydrate-dependent recognition process orients the molecules in a way that brings about specific protein–protein or protein–carbohydrate interactions. As these interactions occur with a unique conformer of the oligosaccharide, the knowledge of the conformation of carbohydrates is important.

A given oligosaccharide can exist in several conformations (Rao, Qasba, Balaji & Chandrasekaran, 1998), and it is the interaction between a unique conformer and a macromolecule that is required to initiate a biological response. Hence, it is essential to have detailed information about all the conformers that are accessible to an oligosaccharide. Given the difficulty in solving carbohydrate structure by X-ray crystallography and as NMR experiments give only time averaged conformations, analysis of oligosaccharide structures by molecular dynamics simulations is a method well suited to determine all the accessible conformations of an oligosaccharide. The objective of such studies is to facilitate the designing of specific sugar analogs that will inhibit carbohydrate–protein interactions and thus lead to the discovery of new therapeutic agents.

Using molecular dynamics simulations, we have investigated the conformational preferences of *N*-linked oligosaccharides—high mannose, complex, and hybrid type oligosaccharides that are the sugar acceptor substrates for glycosyltransferases (reviewed in Qasba, Balaji & Rao, 1997).The conformational data has been correlated a) to the known binding affinities of the oligosaccharide for receptors and b) to various biological processes. These results can be summarized as follows:

1. The MD simulations of several Asn-linked oligosaccharides have provided a wealth of information about their preferred and accessible conformations and have rationalized some of the biochemical and spectroscopic observations.
2. The conformational preferences of the interglycosidic torsion angles in the Asn-linked oligosaccharides are interdependent. Even those saccharide units, which are distant in the primary sequence of the oligosaccharide, may affect the conformational preferences of a disaccharide fragment because of spatial proximity in the oligosaccharide. Significant differences in the conformational preferences of interglycosidic torsion angles are also brought about by the addition/deletion of residues. Hence, the probable conformations of an oligosaccharide cannot be derived from the conformational studies of its constituent di and trisaccharides alone.
3. Asn-linked oligosaccharides are flexible molecules, and the flexibility of the $\alpha1 \rightarrow 3$- and $\alpha1 \rightarrow 6$-linkages in the common pentasaccharide core, compared to the flexibility associated with other linkages, play an important role in determining the overall ‘shape’ of the oligosaccharide

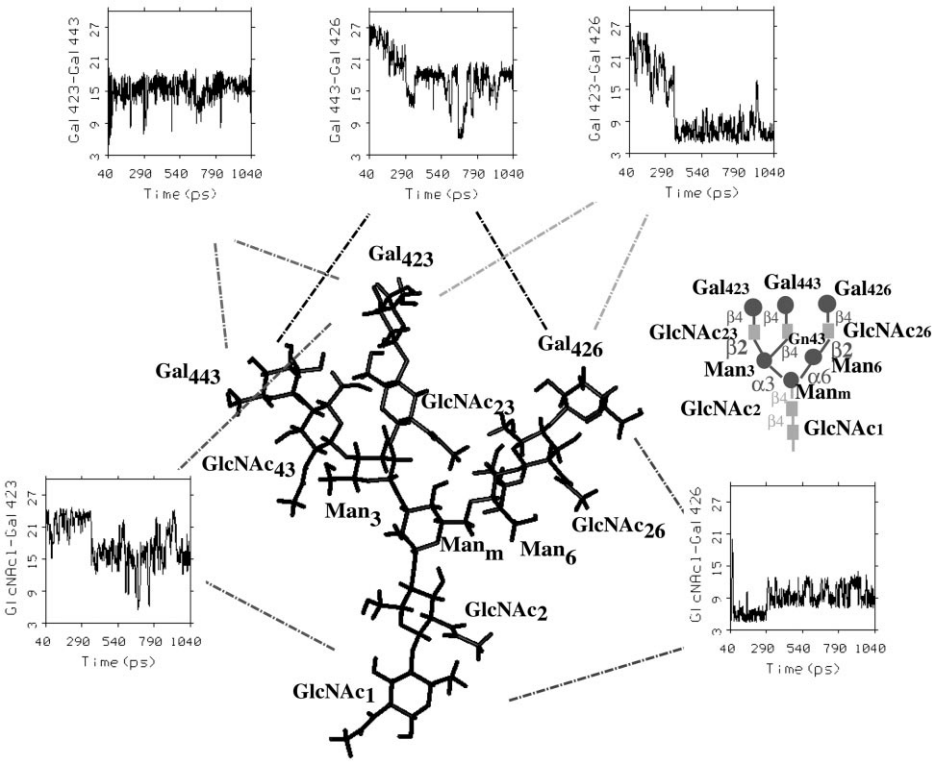

Fig. 4. Variation in distance (Å) between the terminal sugars during 1 ns molecular dynamics simulation of a tri-antennary complex oligosaccharide (sequence shown on the right). Notice the distance between the two terminal galactoses on the 1-3 arm, Gal_{443} and Gal_{423} does not change dramatically over the entire simulation period (top, left corner panel). Also, the distance between galactose on the 1-6-arm (Gal_{426}) and $GlcNAc_1$ on the chitobiose core does not drastically change over the entire simulation period (bottom, right corner panel). On the contrary, the distance between the galactose residues on the 1-6-arm (Gal_{426}) and 1-3 arm (Gal_{426}–Gal_{443} (top, middle panel) or Gal_{426}–Gal_{423} (top, right panel)) does change over the entire simulation period.

(Figs. 4 and 5). Changes in the orientation of the $\alpha 1 \rightarrow 6$-arm in Asn-linked oligosaccharides are brought about not only by changes in χ (Fig. 6) but also by changes in ϕ and ψ(for the same χ (Fig. 7).

4. Processing of Man9GlcNAc2 to Man5GlcNAc2 during the biosynthesis of Asn-linked oligosaccharides is 'conformation driven' and proceeds in a well-defined sequential manner (Balaji et al., 1994). Possible pathways for this processing have been proposed and some of them are in agreement with previous experimental studies. The preferred conformations of Man9GlcNAc2 on the glycoprotein are important in initiating a particular processing pathway.

6. Role of oligosaccharide moiety in glycoprotein folding

N-glycosylation in newly synthesized glycoproteins is important in the folding of the protein. *N*-glycosylated polypeptides entering the lumen of the endoplasmic reticulum are deglucosylated by glucosidases I and II (Fig. 8). In mammalian cells, a lectin-based chaperone system, calnexin, calreticulin and associated factors bind both to monoglucosylated oligosaccharide and the protein moiety and this complex promotes the correct folding of the protein (Fig. 8) (Trombetta & Helenius, 1998; Williams, 1995). To prevent the interaction of the lectin with the folding intermediate, glucosidase II removes the remaining glucose residue, unless the folding intermediate is to be reglucosylated by UDP-glucose:glycoprotein glucosyltransferase (Gluc_T) for recycling. Gluc_T has a unique property in that it glucosylates only the misfolded protein and not the native protein by recognizing two elements in the misfolded protein: the innermost *N*-acetylglucosamine unit of the oligosaccharide and the protein domains exposed in the denatured, but not in the native protein (Sousa & Parodi, 1995). Thus Gluc_T acts as a folding sensor in the system. The glycoprotein stays in the on-and-off cycle until it is correctly folded and no longer requires reglucosylation by Gluc_T. In *Saccharomyces cerevisiae*, specific trimming intermediates of *N*-linked oligosaccharide structure trigger the degradation of the misfolded protein (Jakob, Burda, Roth & Aebi, 1998). The *N*-linked oligosaccharide of the glycoprotein is trimmed by glycosidases to the Man8GlcNAc2 structure (Fig. 9 M8). The properly folded protein (PFP) is exported to the Golgi compartment. If the folding of the protein is not correct the chaperone associated misfolded protein bearing Man8GlcNAc2 structure is recognized by a lectin (Fig. 9 MFP-L) and targeted for export to the cytosol for degradation.

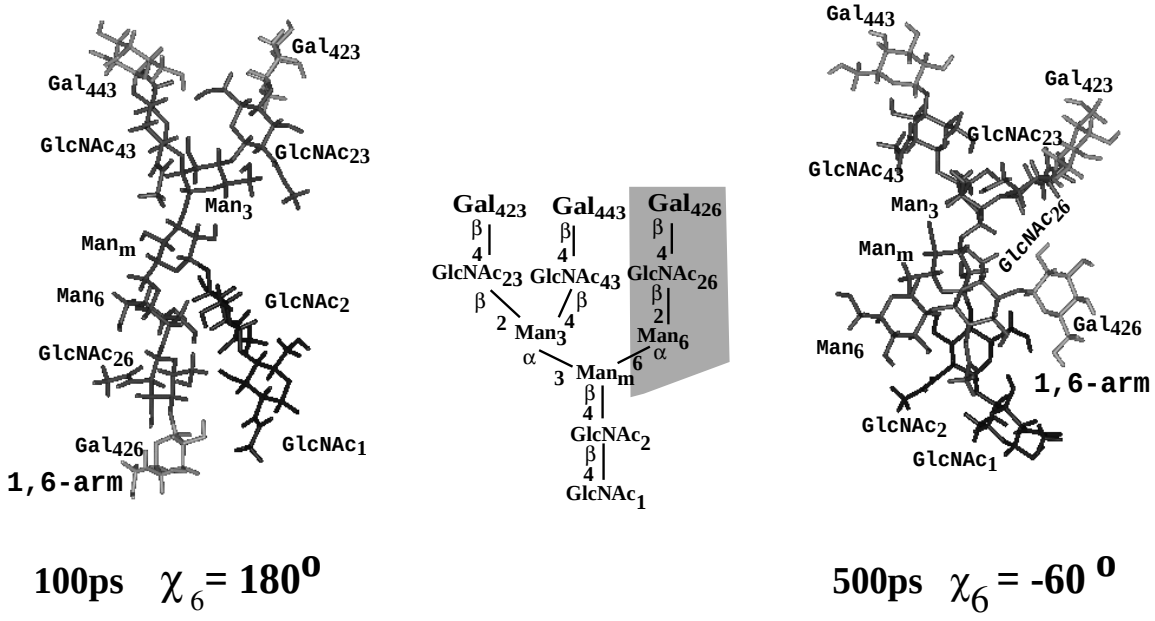

Fig. 5. Snapshots of conformers of tri-antennary oligosaccharide from a 1 ns molecular dynamic simulation. At 100 ps simulation (left) the torsional angle χ of 1-6-arm (shaded portion of the structure shown in the middle) is 180° and the arm is close to chitobiose ($GlcNAc_1$–$GlcNAc_2$). On the contrary, when the torsional angle χ is −60° (right, 500 ps snapshot), the 1-6-arm is at a right angle to chitobiose. (Adapted from Balaji, Qasba & Rao, 1993).

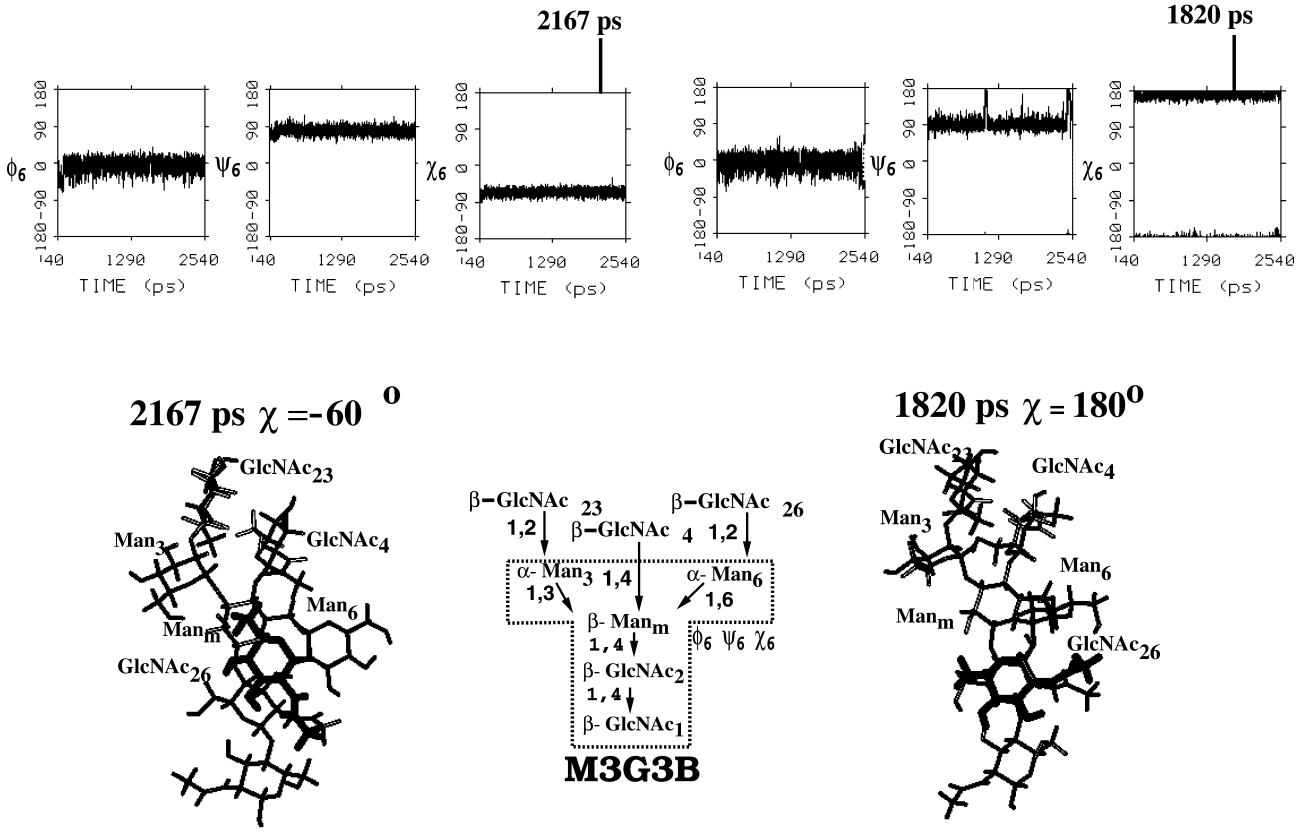

Fig. 6. Changes in the torsional angle χ of the 1-6-linkage change the orientation of the 1-6-arm. Two snapshots of a bisecting bi-antennary oligosaccharide (M3G3B, structure shown in the middle) from a 2.5 ns molecular dynamic simulation show different orientations of the GlcNAc residue on the 1-6-arm ($GlcNAc_{26}$). When χ of the 1-6-linkage is −60° (left, 2167 ps snapshot), the $GlcNAc_{26}$ residue is on the back side, behind the chitobiose core, whereas when the residue is in the front, oriented 90° to the chitobiose core, χ is 180° (right, 1820 ps snapshot). Shown at the top are the fluctuations in ϕ, ψ and χ of the 1-6-linkage over the entire simulation period with two starting confirmations: $\chi = -60°$ (left three panels) and $\chi = 180°$ (right three panels) (adapted from Balaji, Qasba & Rao, 1996).

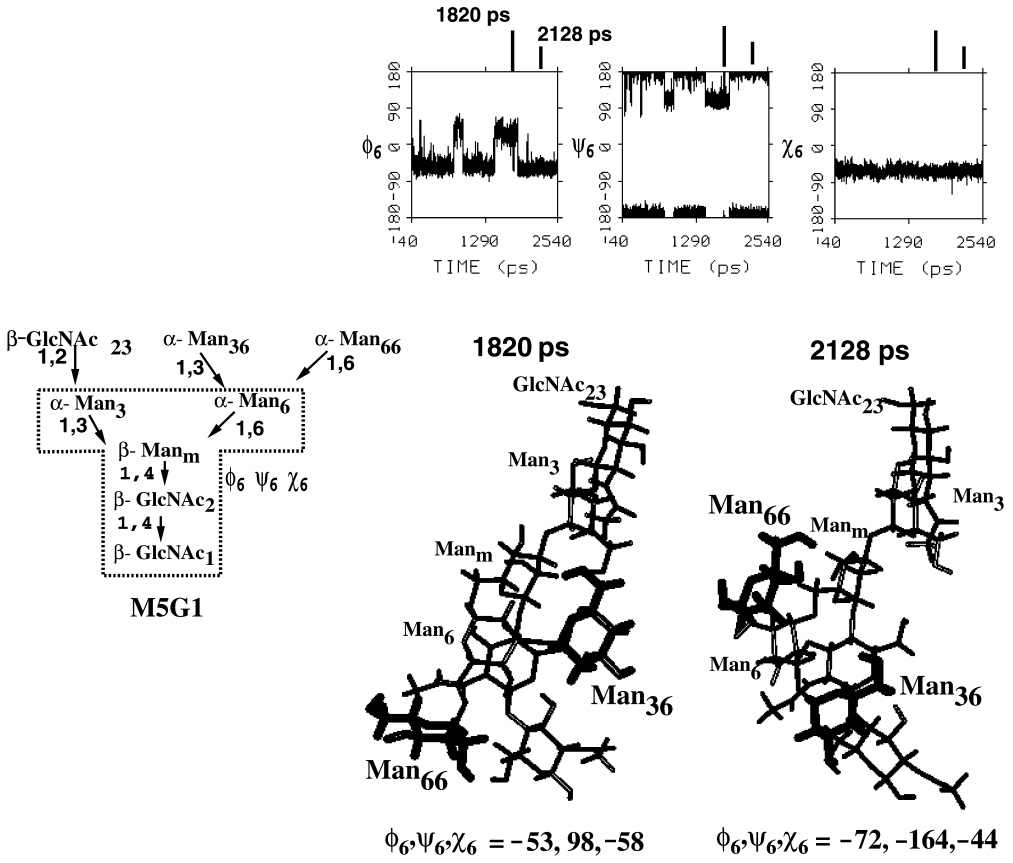

Fig. 7. Changes in the orientation of the 1-6-arm can also be brought about by changing ψ of the 1-6-linkage while keeping χ constant. Two snapshots of a hybrid bi-antennary oligosaccharide (M5G1, structure shown on left) from a 2.5 ns molecular dynamic simulation showing the orientation of mannose residues on the 1-6-arm (Man_{36} and Man_{66}). When ψ of 1-6-linkage is 98° (left, 1820 ps snapshot) the Man_{66} residue is close to the $GlcNAc_1$ residue of chitobiose, and the Man_{36} residue is away and at 90° to chitobiose. On the contrary, when ψ of the 1-6-linkage is −164° (right, 2128 ps snapshot) Man_{36} is close to the chitobiose core and Man_{66} is away and close to the Man_6 residue. Shown in the top three panels are the changes in ϕ, ψ and χ of the 1-6-linkage over the entire simulation period with a starting conformation of $\chi = -60°$ (adapted from Balaji et al., 1996).

7. Defective glycan synthesis and inborn errors

Defects in glycosylation in intact animals have been shown to have serious pathological consequences, and several human diseases have been reported to be the result of faulty glycosylation (reviewed by Varki & Marth, 1995). A defective Golgi-bound enzyme, UDP-GlcNAc:lysosomal enzyme GlcNAc-1-phosphotransferase, which causes a defect in the lysosomal storage of both mucopolysaccharides and glycolipids, results in Inclusion Cell (I-cell) disease where patients show severe psychomotor retardation and death in the first year. The enzyme α1-3fucosyltransferase VII plays an essential role in L-, E-, and P-selectin biosynthesis. These selectins control leukocyte trafficking. The lack of this enzyme results in Leukocyte Adhesion Deficiency Type II (LADII) disease. These patients have an immune deficiency, no recruitment of neutrophils to inflamed tissue, recurrent infections and mental retardation. Recently, certain diseases have been associated with defects in glycosylation and as such, are named Carbohydrate Deficient Glycoprotein Syndrome Type I and Type II. The inactive GlcNAc-transferase II gene has been associated with Carbohydrate-Deficient Glycoprotein Syndrome Type II (CDGS II). The defective mannose specific-glucose independent transporter and defective phosphomannomutase gene (PMM) have been associated with Carbohydrate-Deficient Glycoprotein Syndrome Type I (CDGS I). Clinical features of CDGS I and CDGS II are similar and include psychomotor retardation, cerebellar hyperplasia and severe speech impairment. Inactivation of genes involved in the synthesis of *N*-linked oligosaccharides by transgenic and gene-targeting methodologies have revealed that many of these genes are involved in earlier embryonic development (Metzler, Gertz, Sarkar, Schachter, Schrader & Marth, 1994). Inactivation of mouse GlcNAc-transferase I gene, the *mgat*-I locus, impairs embryogenesis past the embryonic day 9. The mouse embryo lacking this gene dies by the embryonic day 10 (Ioffe & Stanley, 1994). During the development of vulva in *Caenorhabditis elegans*, several mutations have been identified which perturb the vulval invagination. The genes associated with these mutations, SQV-2 and SQV-8, have high sequence similarity with the mammalian β-1,4-galactosyltransferase and glucuronyltransferase, respectively, suggesting that the oligosaccharide structures synthesized by the glycosyltransferases

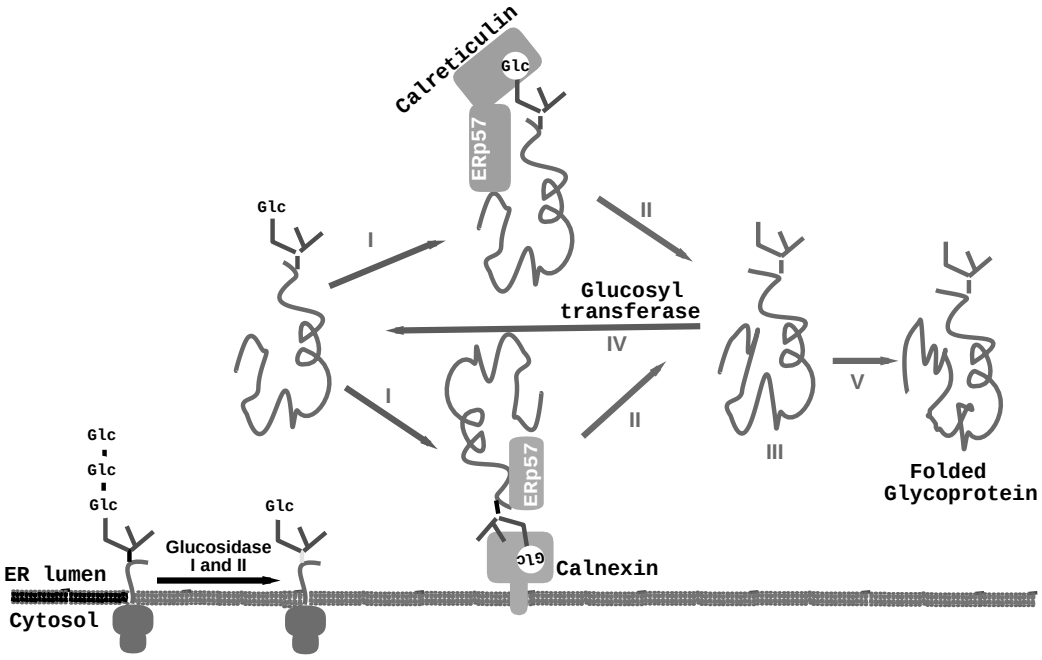

Fig. 8. *N*-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum (ER). The two glucose residues of the *N*-linked core sugar, Glc3Man9GlcNAc2, of the polypeptide chain are removed in the ER by glucosidases I and II. The resulting monoglucosylated glycoproteins bind in the ER to calnexin or calreticulin and are presented to the ERp57 protein. To prevent the interaction of lectin with the folding intermediate, glucosidase II removes the remaining glucose residue. If the folding intermediate does not correctly fold, it is reglucosylated by glucosyltransferase and recycled for the refolding step. The glycoprotein remains in the on-and-off cycle until it is properly folded (adapted from Trombetta & Helenius, 1998).

are crucial during embryonic development (Herman & Horvitz, 1997).

8. Oligosaccharide and protein features recognized in protein–glycoprotein interactions

There is now growing evidence suggesting that in certain instances during protein–glycoprotein interactions, both the oligosaccharide moiety and certain features of the protein are involved in the recognition process. During the glycoprotein folding, as mentioned above, it is the monoglucosylated oligosaccharide moiety together with certain features of the unfolded protein that are recognized by the lectin-based chaperone system to promote the correct folding of the protein (Fig. 8). Several glycosyltransferases have been

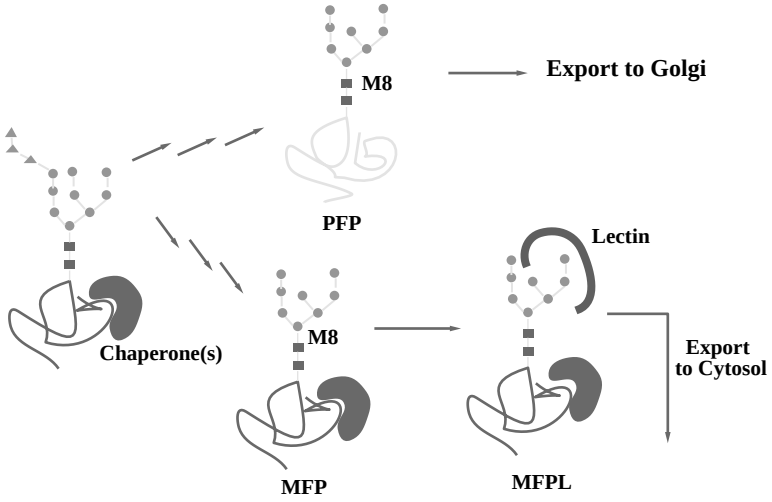

Fig. 9. Role of Man8GlcNAc2 structure in folding of glycoproteins in *Saccharomyces cerevisiae*. With the help of chaperones the glycoproteins are folded in the ER. The glucosidases and mannosidases trim the Glc3Man9GlcNAc2 structure to M8GlcNAc2 glycoprotein. The correctly folded protein (PFP) is exported to the Golgi compartment, while the misfolded protein (MFP) bearing the M8GlcNAc2 structure together with associated chaperones (MFP-L) is exported to the cytosol for degradation (adapted from Jakob, Burda, Roth & Aebi, 1998).

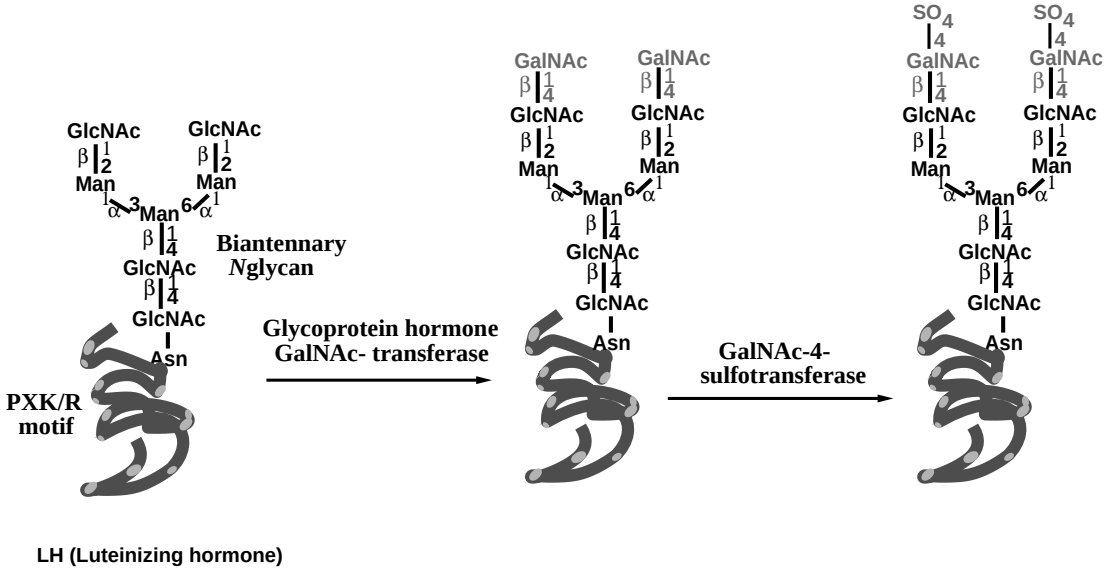

Fig. 10. Oligosaccharide and protein features recognized in protein-glycoprotein interactions: The β1,4-GalNAcT enzyme (glycoprotein hormone GalNAc-transferase) adds GalNAc via a β1,4-linkage to a bi-antennary *N*-glycan of glycoproteins that have specific recognition determinants (gray shaded patches) for the enzyme.

shown to display peptide as well as oligosaccharide specificity. The UDP-glucose:glycoprotein glucosyltransferase that transfers glucose to the misfolded protein for recycling also recognizes both the innermost *N*-acetylglucosamine unit of the oligosaccharide and a denatured domain on the misfolded protein (Fig. 8) (Sousa & Parodi, 1995). The enzyme UDP-GalNAc:glycoprotein hormone *N*-acetylgalactosaminyltransferase adds GalNAc to the oligosaccharide moiety of certain glycoprotein hormones such as lutropin and thyrotropin. This enzyme also recognizes certain features of the peptide chain and the Asn-linked oligosaccharide structure on the protein (Fig. 10) (Manzella et al., 1996). The enzyme β-1,4-galactosyltransferase interacts with the modifier protein α-lactalbumin which changes the sugar acceptor specificity of the enzyme. These interactions between β-1,4-galactosyltransferase and α-lactalbumin occur only in the presence of a sugar or sugar nucleotide (Fig. 11) (reviewed in Qasba et al., 1997). The

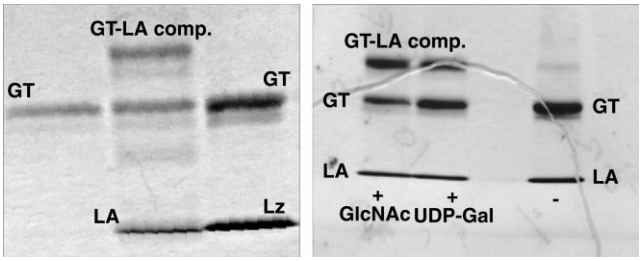

Fig. 11. Sugar or sugar-nucleotide dependent cross-linking of β-1-4-galactosyltransferase (GT) and α-lactalbumin (LA). Polyacrylamide-SDS gel analysis of GT, LA, lysozyme (LZ) and the cross-linked GT and LA (GT-LA complex). The GT-LA complex is formed only with LA, and not with LZ (left). The cross-linking of GT and LA requires either GlcNAc or UDP-Gal (right gel). In the absence of either molecule no cross-linking occurs (extreme right lane).

enzyme UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase, which transfers GlcNAc-phosphate to the oligosaccharide moiety of lysosomal enzymes that are targeted to lysosomes, recognizes a common feature on the lysosomal enzymes as well as on the oligosaccharide moiety (Fig. 12) (Baranski, Cantor & Kornfeld, 1992; Cantor, Baranski & Kornfeld, 1992; Cantor & Kornfeld, 1992). Also, in *Saccharomyces cerevisiae*, as mentioned above, a specific trimming intermediate of *N*-linked oligosaccharide structure is recognized together with a feature in a misfolded protein by a lectin. This complex formation targets the misfolded protein for export to the cytosol for degradation (Fig. 9) (Jakob et al., 1998).

9. Multiple interactive modes between oligosaccharides and proteins

The three-dimensional structures of many protein–carbohydrate complexes have been determined by X-ray crystallography. Analysis of these structures suggests that there are a variety of binding modes between proteins and carbohydrates. Generally oligosaccharides bind with low affinity to the loop regions of the protein which form shallow grooves close to the protein surface. Usually only a terminal sugar residue is involved, and the interaction is limited to a single site on the protein. The high affinity of binding in these cases, results from the clustering of several identical binding sites by formation of protein oligomers. On the contrary, in the sugar transporter proteins monosaccharides bind in the inner cleft of the protein (Quiocho, 1988). The sugar–protein interactions involve multiple hydroxyl residues of the sugar molecule that form the network of hydrogen bonds with both the protein and water molecules. The non-polar

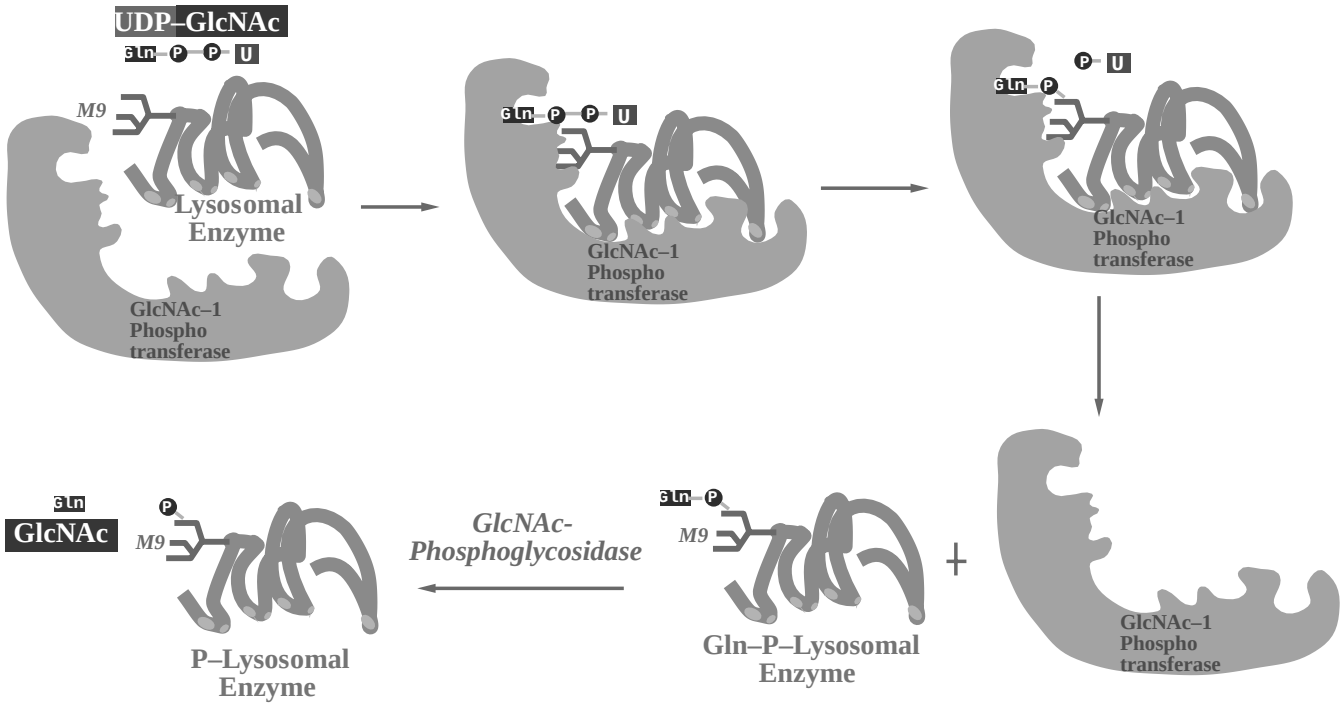

Fig. 12. GlcNAc-1-phosphotransferase enzyme recognizes both the oligosaccharide moiety and certain features on the lysosomal hydrolases that are targeted to lysosomes. In the Golgi apparatus, GlcNAc phosphotransferase enzyme transfers GlcNAc-P- from UDP-GlcNAc to a high mannose *N*-linked oligosacchride of lysosomal hydrolases that carry specific recognition determinants (gray shaded patches) for the enzyme.

interactions between the relatively apolar face of the sugar moiety and the aromatic residues of the protein also contribute significantly towards the stability of the sugar–protein complex. Some of the general sugar–protein binding modes that have emerged from protein–carbohydrate crystal structures are summarized below.

9.1. Multiple sugar bindings in protein–sugar complexes

Calcium-dependent animal lectins, e.g. mannose-binding proteins (MBP) recognize cell-surface oligosaccharides with an intrinsic weak affinity for monosaccharides. However, they bind intensely to multivalent ligands. The polypeptide subunit of these proteins contains carboxy-terminal carbohydrate-binding domains (CRD) (Fig. 13A) and amino-terminal collagenous domains. These proteins associate into trimers (Fig. 13B), which then further assemble into higher multimers forming 'bouquet-like' structures (Fig. 13C). The crystal structure of the CRD domain of MBP in complex with the oligosaccharide ligand (Weis, Drickamer & Hendrickson, 1992) reveals that carbohydrate specificity is determined by a network of coordination and hydrogen bonds that stabilizes the ternary complex of protein, Ca^{2+} and sugar (Fig. 13A). The interactions of the sugar with the CRD occur with the residues in the loop region of the protein and are mediated by many ordered water molecules. In contrast to the legume lectin concanavalin A (Derewenda et al., 1989), *Erythrina corallodendron* lectin (Shaanan, Lis & Sharon, 1991), and *Lathyrus ochrus* isolectin I (Bourne, Roussel, Frey, Rouge, Fontecilla-Camps & Cambillau, 1990), where Ca^{2+} and Mn^{2+} do not directly interact with the sugar but maintain the structural integrity of the binding site (by helping to position the amino acid side chains for sugar binding), the Ca^{2+} in MBP directly ligands to sugar molecule. There is no significant aromatic stacking interactions between mannose and the MBP CRD, and there are fewer van der Waals interactions. The glycosidic bond torsion angles of Man6GlcNAc2 in the CRD complex fall in the same energy minima as those observed in Man9GlcNAc2 in solution. The crystal structure of the trimeric C-type mannose-binding protein (Weis & Drickamer, 1994) reveals that the neck domain, which forms coiled coil α-helices, links the CRD to the collagen-like portion of the intact protein. The structure reveals that there is a 53 Å distance between the carbohydrate-binding sites which is too far apart for a single trimer to bind multivalently to a typical mammalian high-mannose oligosaccharide. To a single MBP trimer an oligosaccharide cannot bind with a high-affinity. Thus it has been suggested by Weis and Drickamer (1994) that the MBP and other collectins, the C-type lectins with collagenous domains, form oligomers that arrange to recognize ligands such as repetitive, very high mannose structures found on fungal and bacterial cell walls (Fig. 13C). This aggregation results in a high affinity binding of the mannose oligosaccharides.

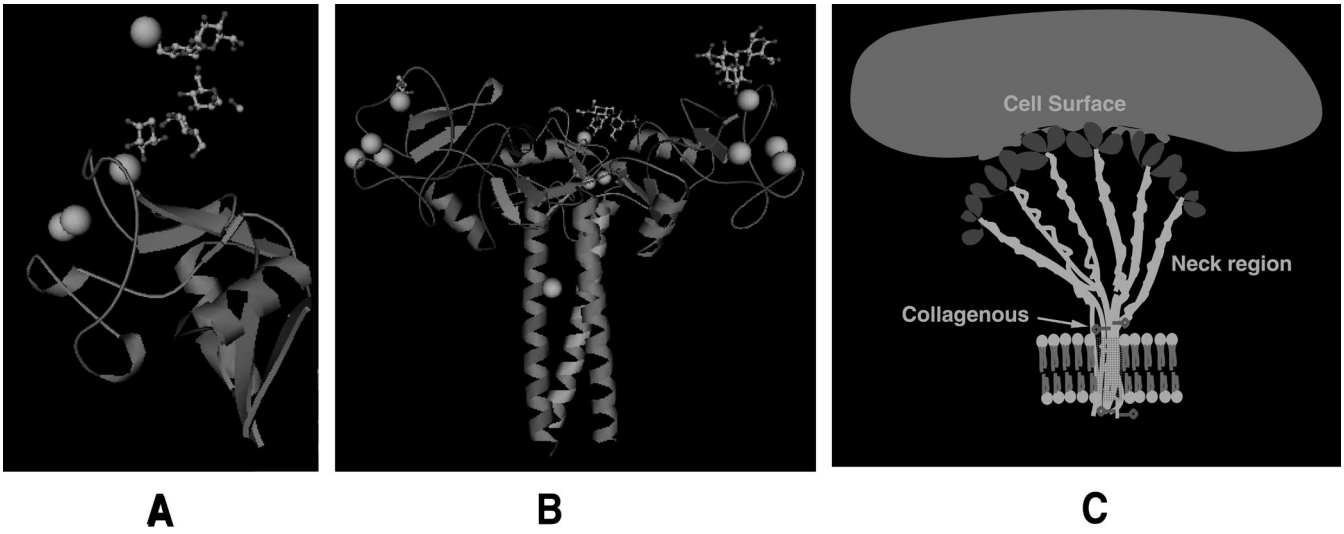

Fig. 13. Multivalent sugar bindings in sugar-protein complexes. (A) Carbohydrate recognition domain (CRD) of mannose binding protein (MBP) in complex with mannose oligosaccharide. The Mannose residue is seen interacting with the loop region of the protein and Ca^{2+}. (B) Trimeric CRD of MBP with neck region in complex with three mannose oligosaccharide structures. The neck region forms coiled coil α-helices. (C) Schematic of the higher multimers of MBP forming 'bouquet-like' structures. The collagenous domain of the assembled structure is embedded in the lipid bilayer and the multivalent binding sites of the structure arrange to recognize high mannose structures of the target surface with high affinity (adapted fromWeis & Drickamer, 1994).

9.2. Cross-linking of proteins by complex oligosaccharides

The structure of wheat germ agglutinin A in complex with a glycopeptide (Wright & Jaeger, 1993) that contains *O*-linked tetrasaccharide, NeuNac-α-(2-3)-Gal-β(1-3)-[NeuNAc-α(2-6)-]-GalNAc-α-1-*O*-Thr, reveals that the oligosaccharide is found in two different conformations in the complex (Fig. 14), an extended conformation in the minor binding site and in a bent conformation in a major binding site. In the minor binding site, the terminal α2,6-NeuNAc residue (Fig. 14, left) is placed in the aromatic-rich site, whereas in the major binding site, the terminal α2,6-NeuNAc and α2,3-NeuNAc residues occupy the sugar specificity sites in the domains on the opposing dimers in the crystal (Fig. 14, right). Such an asymmetric selection of binding sites leads to infinite open-ended arrays of inter-linked lectin molecules and it has been proposed (Wright, 1992) that this asymmetric arrangement of protein binding by an oligosaccharide may serve as model for the interaction of oligosaccharides with complex cell-surface receptors.

Direct cross linking of proteins by an oligosaccharide has been observed in a high resolution three-dimensional structure of three crystal forms of bovine galectin-1 in complex with bi-antennary octasaccharides and asparaginyl nonasaccharide that contain *N*-acetyllactosamine at the non-reducing end of the branches (Fig. 15) (Bourne et al., 1994). The crystal structure reveals infinite chains of lectin dimers cross-linked through *N*-acetyllactosamine units of the oligosaccharides. In each structure the sugar molecules are well defined. The carbohydrate-binding sites of the two, galectin molecules are located at the far ends of the dimers. The high binding specificity for the terminal galactose is mediated by hydrogen bond interactions between the axial 4-hydroxyl group and the side chain residues of His, Asn, and Arg, residues conserved among the galectin family members. The glycosidic torsional angles of the oligosaccharide in the complex fall in the low-energy conformation as observed in solution, except in one of the crystal forms where the α-1-6-linkage differs slightly from the absolute minimum. The main differences in the conformation of the oligosaccharides in the three crystal forms are in the α-1-6- and α-1-3-linkages. In two crystal forms the α-1-6-linkage adopts the *gauche–gauche* conformation, whereas it has a *gauche–trans* conformation in the third crystal form. The ensemble of cross-linked molecules in different crystal forms is the result of the conformational adjustment of α-1-6- and α-1-3-linkages, in part because of the different values for χ angle of the α-1-6-linkage.

9.3. Oligosaccharide acting as a double sided tape

The carbohydrate moieties of proteoglycans, the glycosaminoglycans, bind to growth factors with a high degree of specificity and induce interactions with growth factor receptors, thereby regulating growth factor activity. The basic and acidic fibroblast growth factors, bFGF and aFGF, are closely related members of the FGF family, mitogens which stimulate the growth of a wide spectrum of target cells (Basilico & Moscatelli, 1992). A common feature of the FGF family is the high affinity towards heparin or heparan sulfate proteoglycans and this interaction is required for cell surface receptor (FGFR) binding. These growth factors have a heparin-binding domain that contains an extended collection of positively charged residues. The minimum heparin sequence that has been shown to be required for bFGF to bind to FGFR1 and elicit a maximum

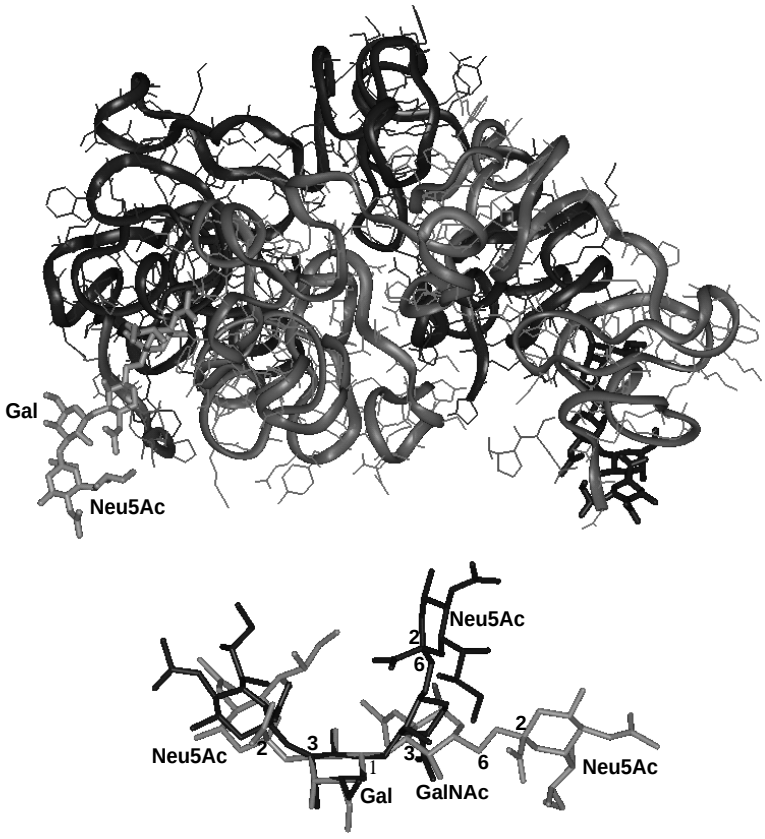

Fig. 14. Complex of Wheat Germ Agglutinin Isolectin I with sialoglycopeptide (upper panel) (Wright, 1992). The two glycopeptide molecules, each containing an *O*-linked tetrasaccharide, NeuNAc-α-(2-3)-Gal-β(1-3)-[NeuNAc-α(2-6)-]-GalNAc-α-1-*O*-, are bound in the complex, one at the minor binding site (left side) and the other at the major binding site (right side), and differ in conformation. Superposition of the oligosaccharides from the two binding sites (lower panel) shows that the tetrasaccharide in the minor binding site is in an extended conformation (gray colored), while the tetrasaccharide in the major binding site is in a bend conformation (black). The bend conformation results from a change in ψ torsional angle of the Gal-β(1-3)-GalNAc-linkage (adapted from Wright, 1992).

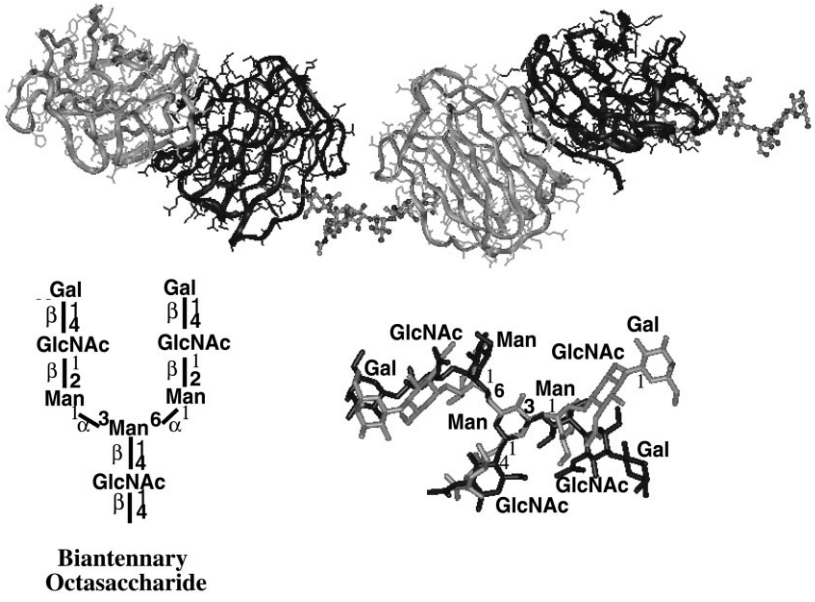

Fig. 15. Cross-linking of Galectin-1 dimers by a bi-antennary octasaccharide. The two carbohydrate binding sites are located at the far ends of the dimers (upper panel). Superposition of the two oligosaccharides from the complex (lower panel) shows that the oligosaccharides at the two binding sites adopt different conformations because of changes in the torsional angles of the α-1-6- and α-1-3-linkages (adapted from Bourne et al., 1994).

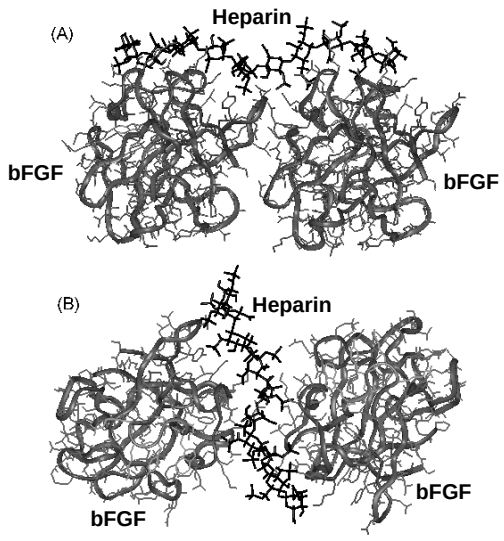

Fig. 16. Heparin mediated dimerization of basic fibroblast growth factor (bFGF). (A) A proposed model of the *cis* dimer of bFGF in complex with a heparin decasaccharide (Lam et al., 1998). (B) A *trans*-dimer of bFGF in complex with a heparin decasaccharide, a model based on the crystal structure of the *trans*-dimer of aFGF in complex with a heparin decasaccharide as described by DiGabriele et al., 1998).

biological response consists of an octa or a decasaccharide structure (Ishihara, Tyrrell, Stauber, Brown, Cousens & Stack, 1993; Yayon, Klagsburn, Esko, Leder & Ornitz, 1991). Both by site-directed mutagenesis and by crystal structure analysis of the complex between heparin and bFGF (Faham, Hileman, Fromm, Linhardt & Rees, 1996) or aFGF (DiGabriele et al., 1998), the residues involved in the heparin binding site have been identified. The structures of bFGF with tetra and hexasaccharides of heparin described by Faham et al. (1996) and, recently, the structure of aFGF in complex with a decasaccharide molecule described by DiGabriele et al. (1998) have further demonstrated the precise location and interactions of the heparin chain in the FGF binding surface. Two additional binding sites on bFGF for small heparin analogs have been localized by Ornitz, Herr, Nilsson, Westman, Svahn & Waksman (1995) in the crystal structure of the complex between bFGF and non-sulfated tri and tetrasaccharides. The heparin binding domain in both bFGF and aFGF, which consists mainly of the side chains of the surface loops makes ionic contacts and hydrogen bonds with sulfate or carboxylate groups of heparin.

Several mechanisms have been proposed by which FGF together with heparin induces the activation of FGFR leading to cell proliferation (Springer et al., 1994). Heparin binding to FGF induces the dimerization of the growth factor (Pantoliano et al., 1994), but the nature of the dimer is still under debate (Herr, Ornitz, Sasisekharan, Venkataraman & Waksman, 1997; Moy et al., 1997; Ornitz et al., 1995). In the structure of aFGF with sucrose octasulfate (SOS), there is one monomer in the asymmetric unit of the crystal. In addition, in the structures of bFGF with tetra and hexasaccharides (Faham et al., 1996), there is only one molecule in the asymmetric unit of the crystal, a situation similar to what has been observed in the crystal structure of bFGF with non-sulfated tri and tetrasaccharides (Ornitz et al., 1995), even though the heparin binding sites in the two structures are different. Binding of small oligosaccharides to bFGF has been studied either by dynamic light scattering and NMR (Moy et al., 1997) or by analytical ultracentrifugation (Herr et al., 1997). It has been concluded that sugars as small as tetrasaccharides or SOS induce dimerization of FGF and the active dimer induced by the heparin molecules with a chain length of eight or more monomer sugar units consists of two molecules of FGF bound to the oligosaccharide on the same face of the chain, a *cis*-dimer (Ornitz et al., 1995). Using molecular modeling method, we have also proposed a *cis*-dimer model for bFGF with heparin decasaccharide (Lam, Rao & Qasba, 1998) (Fig. 16A) that satisfies the mutational data and the requirement for the FGF receptor to bind heparin in the complex. Recently, DiGabriele et al. (1998) described the structures of aFGF with a decasaccharide where four dimers are present in the asymmetric unit of an orthorhombic crystal, whereas in a hexagonal crystal form one dimer is present in the asymmetric unit. The aFGF molecules in the dimers are in *trans* configuration (Fig. 16B) where one protomer is bound oppositely to the aFGF-dimer mate; thus heparin acts as a double sided tape. This was the first structural evidence for the heparin-induced dimerization of aFGF.

We have previously built a 3D model of fibroblast growth factor receptor 1 (FGFR1) based on the homology of the D(II) and D(III) domains of the receptor to the variable region of the IgG light and heavy chains (Lam et al., 1998). The modeled FGFRI was docked to the *cis*-dimer of bFGF with heparin, and the docked complex (Fig. 17, left) was shown to satisfy already known mutational data. Further, the heparin-binding site of FGFR1 was interacting with heparin in the complex. We have presently docked the modeled FGFR1 receptor to the modeled *trans*-dimer of bFGF–heparin complex in a way that satisfies the criteria for the interactions with the heparin molecule and with the primary and secondary binding sites on the bFGF molecule. Two possible models (Fig. 17A and B) can satisfy these criteria. In model A, each FGFR1 molecule interacts with the primary and secondary binding sites located on the same bFGF molecule. In model B, each FGFR1 molecule interacts with the primary site on one molecule of bFGF and the secondary site of the other bFGF molecule. This model results in cross-linking of the two, bFGF molecules by each FGFR1 molecule. Compared with our previously published data (Lam et al., 1998), these molecular modeling and docking studies show that the complex of FGFR1 with the *trans*-dimer of bFGF is energetically unfavorable. The *cis*-modeled structure is a more favorable complex and is consistent with the binding mechanism of FGF to its receptor, the receptor dimerizaton, and the reported site-specific

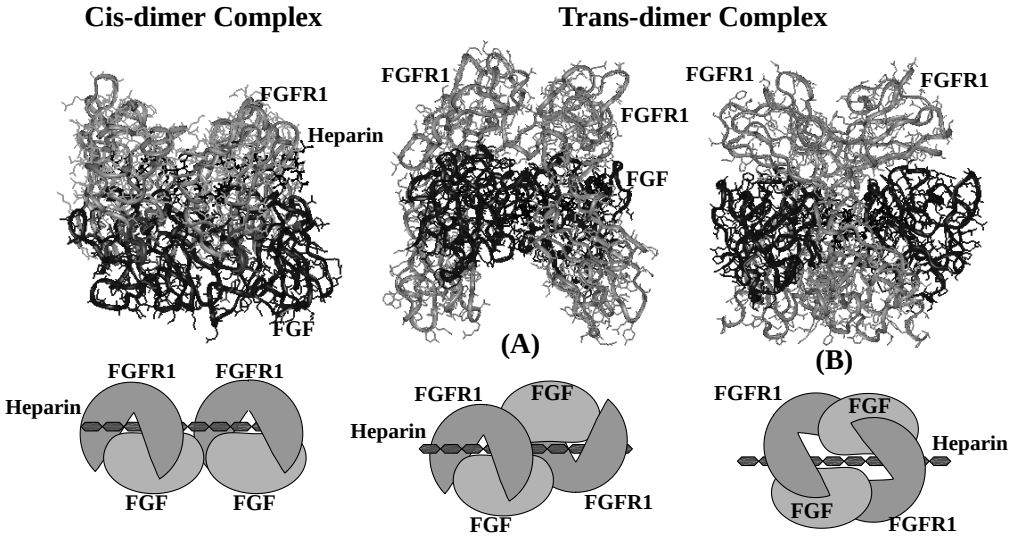

Fig. 17. Modeled complexes of FGFR1 with *cis*- and *trans*-dimers of bFGF and a heparin molecule. Pictures on the top are FGFR1 dimers the docked to the *cis*-dimer of bFGF and heparin molecule (left), and the two possible models of the complex of FGFR1 with the *trans*-dimer of bFGF and heparin ((A) and (B)). Lower panels show the schematic orientation of the interacting molecules.

mutagenesis and biochemical cross-linking data. The crystal structure of the complex between FGFR1, FGF, and heparin will definitively establish the modes of binding between these molecules. Nevertheless, the recently described structure of aFGF with a heparin decasaccharide by DiGabriele et al. (1998) and our molecular modeling approach, support the hypothesis that the sugars orient the two protein molecules in a way that bring about specific protein–protein or protein–carbohydrate interactions.

10. Conclusions

1. Sugars utilize hydroxyl groups to form hydrogen bonds with a protein either directly or indirectly through water molecules. As specific orientations of non-polar CH groups in a sugar residue create a hydrophobic patch, sugars can also make hydrophobic interactions with proteins.
2. Generally, one or two metal ions bind near a sugar binding site of a protein and help to position the peptide element near the binding site. Sometimes, sugar hydroxyls may directly coordinate with the metal ion as in mannose binding protein.
3. Sugars generally bind weakly to proteins in shallow grooves close to the surface, interacting with the flexible loop region of the protein. The binding sites are not highly discriminative but exhibit multiple specificities.
4. Oligosaccharides upon binding may induce conformational changes in the binding pocket, particularly when they bind in the loop region of the protein.
5. The conformational flexibility of oligosaccharides, particularly around the 1,3- and 1,6-arms, brings about significant changes in ligand shape, which results in proper orientation of the side groups and favorable interactions with the binding residues of the proteins.
6. Binding of each antenna of a multi-antennary oligosaccharide brings together or cross-links two protein molecules in a specific way.
7. Oligosaccharides, in particular glycosaminoglycans, may act as a double sided tape, in that they bring together or cross-link two protein molecules in a specific orientation, e.g. heparin bridges the two FGF molecules in a specific orientation and the resulting complex induces FGFR1 dimerization.
8. In the oligosaccharide recognition process, sugars induce specific geometries of interaction between protein molecules that result in a specific biological response.

Acknowledgements

I thank my former associates, Drs V.S.R. Rao, P.V. Balaji and K. Lam, and the present associates Drs F. Bizik, B. Ramakrishnan and E. Boeggeman who have contributed to my present knowledge and understanding of the carbohydrate structure and function. I also acknowledge the Frederick Biomedical Super Computing Center of the FCRDC, NCI, for allocation of computing time and staff support.

References

Almeida, R., Amado, M., David, L., Levery, S. B., Holmes, E. H., Merkx, G., van Kessel, A. G., Rygaard, E., Hassan, H., Bennett, E., & Clausen, H. (1997). A family of human beta4-galactosyltransferases. Cloning and expression of two novel UDP-galactose:beta-n-acetylglucosamine beta1, 4-galactosyltransferases, beta4Gal-T2 and beta4Gal-T3 [published erratum appears in 1998 Jul 17, *Journal of Biological*

Chemistry, 273(29), 18674] *Journal of Biological Chemistry*, 272, 31979–31991.

Balaji, P. V., Qasba, P. K., & Rao, V. S. R. (1993). Molecular dynamics simulations of asialoglycoprotein receptor ligands. *Biochemistry*, *32*, 12599–12611.

Balaji, P. V., Qasba, P. K., & Rao, V. S. R. (1994). Molecular dynamics simulations of high-mannose oligosaccharides. *Glycobiology*, *4*, 497–515.

Balaji, P. V., Qasba, P. K., & Rao, V. S. R. (1996). Molecular dynamics simulations of hybrid and complex type oligosaccharides. *International Journal of Biological Macromolecules*, *18*, 101–114.

Baranski, T. J., Cantor, A. B., & Kornfeld, S. (1992). Lysosomal enzyme phosphorylation. I. Protein recognition determinants in both lobes of procathepsin D mediate its interaction with UDP-GlcNAc:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase. *Journal of Biological Chemistry*, *267*, 23342–23348.

Basilico, C., & Moscatelli, D. (1992). The FGF family of growth factors and oncogenes. *Adv. Cancer Res.*, *59*, 115–165.

Bischoff, J., Liscum, L., & Kornfeld, R. (1986). The use of 1-deoxymannojirimycin to evaluate the role of various a-mannosidases in oligosaccharide processing in intact cells. *Journal of Biological Chemistry*, *261*, 4766–4774.

Bourne, Y., Bolgiano, B., Liao, D. I., Strecker, G., Cantau, P., Herzberg, O., Feizi, T., & Cambillau, C. (1994). Crosslinking of mammalian lectin (galectin-1) by complex bi-antennary saccharides. *Natural Structure and Biology*, *1*, 863–870.

Bourne, Y., Roussel, A., Frey, M., Rouge, P., Fontecilla-Camps, J. C., & Cambillau, C. (1990). Three-dimensional structures of complexes of *Lathyrus ochrus* isolectin I with glucose and mannose: fine specificity of the monosaccharide-binding site. *Proteins*, *8*, 365–376.

Brockhausen, I., & Schachter, H. (1997). Glycosyltransferases involved in *N*- and *O*-glycan biosynthesis. In S. Gabius (Ed.), *Glyco-sciences: status and perspectives*, (pp. 79–113). London: Chapman & Hall.

Cantor, A. B., Baranski, T. J., & Kornfeld, S. (1992). Lysosomal enzyme phosphorylation. II. Protein recognition determinants in either lobe of procathepsin D are sufficient for phosphorylation of both the amino and carboxyl lobe oligosaccharides. *Journal of Biological Chemistry*, *267*, 23349–23356.

Cantor, A. B., & Kornfeld, S. (1992). Phosphorylation of Asn-linked oligosaccharides located at novel sites on the lysosomal enzyme cathepsin D. *Journal of Biological Chemistry*, *267*, 23357–23363.

Cohen, R. E., & Ballou, C. E. (1980). Linkage and sequence analysis of mannose-rich glycoprotein core oligosaccharides by proton nuclear magnetic resonance spectroscopy. *Biochemistry*, *19*, 4345–4358.

Collier, E., Carpentier, J. L., Beitz, L., Carol, H., Taylor, S. I., & Gorden, P. (1993). Specific glycosylation site mutations of the insulin receptor alpha subunit impair intracellular transport. *Biochemistry*, *32*, 7818–7823.

Dahms, N. M., & Hart, G. W. (1986). Influence of quaternary structure on glycosylation: differential subunit association affects the site-specific glycosylation of the common beta-chain from Mac-1 and LFA-1. *Journal of Biological Chemistry*, *261*, 13186–13196.

Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J., Papiz, M. Z., Wan, T., & Campbell, J. (1989). The structure of the saccharide-binding site of concanavalin A. *EMBO Journal*, *8*, 2189–2193.

DiGabriele, A. D., Lax, I., Chen, D. I., Svahn, C. M., Jayes, M., Schlessinger, J., & Hendrickson, W. A. (1998). Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature*, *393*, 812–817.

Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J., & Rees, D. C. (1996). Heparin structure and interactions with basic fibroblast growth factor. *Science*, *271* (5252), 1116–1120.

Fast, D. G., Jamieson, J. C., & McCaffrey, J. (1993). The role of the carbohydrate chains of Gal-β1,4-GlcNAc-β2,6-Sialyltransferase for enzymatic activity. *Biochimica Biophysica Acta*, *1202*, 325–340.

Green, E. D., & Baenziger, J. U. (1988). Asparagine-linked oligosaccharides on lutropin, follitropin, and thyrotropin. I. Structural elucidation of the sulfated and sialylated oligosaccharides on bovine, ovine, and human pituitary glycoprotein hormones. *Journal of Biological Chemistry*, *263*, 25–35.

Haltiwanger, R. S., Kelly, W. G., Roquemore, E. P., Blomberg, M. A., Dong, L. Y., Kreppel, L., Chou, T. Y., & Hart, G. W. (1992). Glycosylation of nuclear and cytoplasmic proteins is ubiquitous and dynamic. *Biochemical Society Transactions*, *20*, 264–269.

Harris, R. J., & Spellman, M. W. (1993). *O*-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, *3*, 219–224.

Herman, T., & Horvitz, H. R. (1997). Mutations that perturb vulval invagenation in *C. elegans*. *Cold Spring Harbor Symposium on Quantitative Biology*, *LXII*, 353–360.

Herr, A. B., Ornitz, D. M., Sasisekharan, R., Venkataraman, G., & Waksman, G. (1997). Heparin-induced self-association of fibroblast growth factor-2. Evidence for two oligomerization processes. *Journal of Biological Chemistry*, *272*, 16382–16389.

Ioffe, E., & Stanley, P. (1994). Mice lacking *N*-acetylglucosaminyltransferase I activity die at mid- gestation, revealing an essential role for complex or hybrid *N*-linked carbohydrates. *Proceedings of the National Academy of Science USA*, *91*, 728–732.

Ishihara, M., Tyrrell, D. J., Stauber, G. B., Brown, S., Cousens, L. S., & Stack, R. J. (1993). Preparation of affinity-fractionated, heparin-derived oligosaccharides and their effects on selected biological activities mediated by basic fibroblast growth factor. *Journal of Biological Chemistry*, *268* (7), 4675–4683.

Jakob, C. A., Burda, P., Roth, J., & Aebi, M. (1998). Degradation of misfolded endoplasmic reticulum glycoproteins in *Saccharomyces cerevisiae* is determined by a specific oligosaccharide structure. *Journal of Biological Chemistry*, *142*, 1223–1233.

Kornfeld, R., & Kornfeld, S. (1985). Assembly of asparagine linked oligosaccharides. *Annual Reviews in Biochemistry*, *54*, 631–664.

Lam, K., Rao, V. S. R., & Qasba, P. K. (1998). Molecular modeling studies on binding of bFGF to heparin and its receptor FGFR1. *Journal of Biomolecular Structure and Dynamics*, *15*, 1009–1027.

Lo, N.-W., Shaper, J. H., Pevsner, J., & Shaper, N. L. (1998). The expanding β4-galactosyltransferase gene family: messages from the databank. *Glycobiology*, *8*, 517–526.

Lubas, W. A., & Spiro, R. G. (1987). Golgi endo-α-D-mannosidase from rat liver, a novel *N*-linked carbohydrate unit processing enzyme. *Journal of Biological Chemistry*, *262*, 3775–3781.

Lubas, W. A., & Spiro, R. G. (1988). Evaluation of the role of rat liver Golgi endo-α-D-mannosidase in processing *N*-linked oligosaccharides. *Journal of Biological Chemistry*, *263*, 3990–3998.

Manzella, S. M., Hooper, L. V., & Baenzinger, J. U. (1996). Oligosaccharide containing b1,4-linked *N*-acetylgalactosamine, a paradigm for protein-specific glycosylation. *Journal of Biological Chemistry*, *271*, 12117–12120.

Metzler, M., Gertz, A., Sarkar, M., Schachter, H., Schrader, J. W., & Marth, J. D. (1994). Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *EMBO Journal*, *13*, 2056–2065.

Moss, L., Prakobphol, A., Wiedmann, T.-W., Fisher, S. J., & Damsky, C. H. (1994). Glycosylation of human trophoblast integrins is stage and cell-type specific. *Glycobiology*, *4*, 567–575.

Moy, F. J., Safran, M., Seddon, A. P., Kitchen, D., Bohlen, P., Aviezer, D., Yayon, A., & Powers, R. (1997). Properly oriented heparin-decasaccharide-induced dimers are the biologically active form of basic fibroblast growth factor. *Biochemistry*, *36* (16), 4782–4791.

Narimatsu, H., Sinha, S., Brew, K., Okayama, H., & Qasba, P. K. (1986). Cloning and sequencing of cDNA of bovine *N*-acetylglucosamine (β1-4)galactosyltransferase. *Proceedings of the National Academy of Science USA*, *83*, 4720–4724.

Ornitz, D. M., Herr, A. B., Nilsson, M., Westman, J., Svahn, C. M., & Waksman, G. (1995). FGF binding and FGF receptor activation by synthetic heparin-derived di- and trisaccharides. *Science*, *268*, 432–436.

Pantoliano, M. W., Horlick, R. A., Springer, B. A., VanDyk, D. E., Tobery,

T., Wetmore, D. R., Lear, J. D., Nahapetian, A. T., Bradley, J. D., & Sisk, W. P. (1994). Multivalent ligand-receptor binding interactions in the fibroblast growth factor system produce a cooperative growth factor and heparin mechanism for receptor dimerization. *Biochemistry, 33* (34), 10229–10248.

Qasba, P. K., Balaji, P. V., & Rao, V. S. R. (1997). Conformational analysis of Asn-linked oligosaccharides: implications in biological processes. *Journal of Molecular Structure (Theochem), 395-396*, 333–360.

Qasba, P. K., & Kumar, S. (1997). Molecular divergence of lysozyme and α-lactalbumin. *Critical Reviews in Biochemistry and Molecular Biology*, *32*, 255–306.

Quiocho, F. A. (1988). Molecular features and basic understanding of protein-carbohydrate interactions: the arabinose-binding protein-sugar complex. *Current Topics in Microbiology and Immunology*, *139*, 137–148.

Rao, V. S. R., Qasba, P. K., Balaji, P. V., & Chandrasekaran, R. (1998). *Conformation of carbohydrates*, Australia: Harwood.

Shaanan, B., Lis, H., & Sharon, N. (1991). Structure of a legume lectin with an ordered *N*-linked carbohydrate in complex with lactose. *Science*, *254*, 862–866.

Shaper, N. L., Shaper, J. H., Meuth, J. L., Fox, J. L., Chang, H., Kirsch, I. R., & Hollis, G. F. (1986). Bovine galactosyltransferase: identification of a clone by direct immunological screening of a cDNA expression library. *Proceedings of the National Academy of Science USA*, *83*, 1573–1577.

Snider, M. D. (1984). Biosynthesis of glycoproteins: formation of *N*-linked oligosaccharides. In P. W. Ginsburg (Ed.), *Biology of Carbohydrates*, (pp. 163–198). New York: Wiley.

Sousa, M., & Parodi, A. J. (1995). The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO Journal*, *14*, 4196–4203.

Springer, B. A., Pantoliano, M. W., Barbera, F. A., Gunyuzlu, P. L., Thompson, L. D., Herblin, W. F., Rosenfeld, S. A., & Book, G. W. (1994). Identification and concerted function of two receptor binding surfaces on basic fibroblast growth factor required for mitogenesis. *Journal of Biological Chemistry*, *269* (43), 26879–26884.

Trombetta, E. S., & Helenius, A. (1998). Lectins as chaperones in glycoprotein folding. *Current Opinion in Structural Biology*, *8*, 587–592.

Varki, A., & Marth, J. (1995). Oligosaccharide in vertebrate development. *Developmental Biology*, *6*, 127–138.

Vliegenthart, J. F. G., & Casset, F. (1998). Novel forms of protein glycosylation. *Current Opinions in Structural Biology*, *8*, 565–571.

Weis, W. I., & Drickamer, K. (1994). Trimeric structure of a C-type mannose-binding protein. *Structure*, *2*, 1227–1240.

Weis, W. I., Drickamer, K., & Hendrickson, W. A. (1992). Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature*, *360*, 127–134.

Weng, S., & Spiro, R. G. (1993). Demonstration that a kifunensine-resistant α-mannosidase with a unique processing action on *N*-linked oligosaccharides occurs in rat liver endoplasmic reticulum and various cultured cells. *Journal of Biological Chemistry*, *268*, 25656–25663.

Williams, A. F., Parekh, R. B., Wing, D. R., Willis, A. C., Barclay, A. N., Dalchau, R., Fabre, J. W., Dwek, R. A., & Rademacher, T. W. (1993). Comparative analysis of the *N*-glycans of rat, mouse and human Thy-1. Site-specific oligosaccharide patterns of neural Thy-1, a member of the immunoglobulin superfamily. *Glycobiology*, *3*, 339–348.

Williams, D. B. (1995). Calnexin: a molecular chaperone with a taste for carbohydrate. *Biochemistry and Cell Biology*, *73*, 123–132.

Wright, C. S. (1992). Crystal structure of a wheat germ agglutinin/glycophorin-sialoglycopeptide receptor complex. Structural basis for cooperative lectin-cell binding. *Journal of Biological Chemistry*, *267*, 14345–14352.

Wright, C. S., & Jaeger, J. (1993). Crystallographic refinement and structure analysis of the complex of wheat germ agglutination with a bivalent sialoglycopeptide from glycophorin A. *J. Mol. Biol.*, *232*, 620–638.

Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., & Ornitz, D. M. (1991). Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell*, *64* (4), 841–848.

Yuan, Y. P., Schultz, J., Mlodzik, M., & Bork, P. (1997). Secreted fringe-like signaling molecules may be glycosyltransferases [letter]. *Cell*, *88*, 9–11.

Zamze, S. E., Wooten, E. W., Ashford, D. A., Ferguson, M. A. J., Dwek, R. A., & Rademacher, T. W. (1990). Characterization of the asparagine-linked oligosaccharides from *Trypanosoma brucei* type-I variant surface glycoproteins. *European Journal of Biochemistry*, *187*, 657–663.

Zhang, Y., & Dahms, N. M. (1993). Site directed removal of *N*-glycosylation sites in the bovine cation-dependent mannose 6-phosphate receptor—effects on ligand binding, intracellular targeting and association with binding immunoglobulin protein. *Biochemical Journal*, *295*, 841–848.