8.2 Glycoproteins: Properties

Valentin Wittmann

Fachbereich Chemie, Universität Konstanz, 78457 Konstanz, Germany

mail@valentin-wittmann.de

1	Introduction	1772
2	Modulatory and Structural Roles of Glycans	1773
2.1	Modulation of Physicochemical Properties	1773
2.2	Protective and Stabilizing Functions	1775
2.3	Modulation of Biological Activity	1776
2.4	Influence on Peptide Secondary Structure	1777
3	Involvement of Glycans in Recognition Events	1780
3.1	Carbohydrate Recognition by Blood Group Antibodies	1781
3.2	Carbohydrate-Modifying Enzymes	1782
3.3	Carbohydrate-Lectin Interactions	1785
3.3.1	Classification of Lectins	1785
3.3.2	Lectin Control of Protein Folding	1786
3.3.3	Clearance and Targeting of Glycoproteins	1787
3.3.4	Leukocyte Trafficking	1788

Abstract

This chapter focuses on the biological roles of the glycans contained in glycoproteins. Today we know there is no unifying function for the carbohydrates present in glycoproteins. They rather span the complete spectrum from being obviously unimportant to being crucial for the survival of an organism. In a crude scheme, their biological functions can be classified into two groups. On one hand, the carbohydrates can modify intrinsic properties of a protein by altering its size, charge, solubility, accessibility, structure, or dynamic properties. On the other hand, the glycans themselves may be specifically recognized by carbohydrate-binding proteins and thus participate in adhesion processes and signal transduction. A selection of some of these processes, that are well characterized, will be highlighted.

Keywords

Glycoproteins; Glycopeptides; Antifreeze glycoproteins; Protein stability; Protein activity; Glycopeptide conformation; Blood group antigens; Glycosyltransferases; Glycosidases; Protein folding

Abbreviations

AFGPs	antifreeze glycoproteins
CD	circular dichroism
CHO	Chinese hamster ovary
CRD	carbohydrate recognition domain
DAF	decay accelerating factor
EGF	epidermal growth factor
ER	endoplasmic reticulum
FDA	Food and Drug Administration
FRET	fluorescence resonance energy transfer
NOE	nuclear Overhauser enhancement
tPA	tissue plasminogen activator

1 Introduction

Whereas the last chapter provided an overview of the structures of glycoproteins and glycopeptides occurring in nature, this chapter focuses on the biological roles of their carbohydrate units. However, the frequently asked question “What is the function of glycosylation?” is actually as unreasonable as would be the question “What is the function of proteins?” Today we know there is no unifying function for the carbohydrates present in glycoproteins. They rather span the complete spectrum from being relatively unimportant to being crucial for the survival of an organism. Moreover, the same glycan may have different functions at different locations on a given protein, or in different cells or developmental stages of an organism. Thus, each glycoprotein has to be studied individually in order to unravel the roles of its glycans. The aim of this chapter is to indicate some general principles of protein glycosylation.

Diverse approaches are being employed in order to uncover the roles of carbohydrates contained in glycoconjugates [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18]. They include the localization of specific glycans using lectins or antibodies, the modification of glycans by glycosyltransferases and glycosidases, and the use of inhibitors of glycan biosynthesis or processing. Natural or synthetic ligands can be used to identify specific carbohydrate receptors. The methods of molecular biology provide a powerful tool to study glycan function. Thus, it is possible to generate cell mutants with altered glycosyltransferase expressions. Alternatively, recombinant glycoproteins can be expressed in different cells with different glycosylation properties. Chinese hamster ovary (CHO) cells, for example, are not able to generate sialic acid α 2–6 linkages and galactose α 1–3 linkages and bacteria like *Escherichia coli* produce completely non-glycosylated proteins. Recently, the chemical synthesis of pure glycoforms of glycoproteins or glycoprotein mimetics has become feasible providing a powerful tool for structure-activity relationship studies [19]. Metabolic engineering is another technique developed by chemists that gives access to glycoproteins containing functional groups that are selectively addressable by chemical ligation reactions [20,21]. Carbohydrate chips are used to probe carbohydrate recognition by proteins and other binding partners in a high-throughput manner [14,22,23,24,25].

In a crude scheme, the biological functions of glycans can be classified into two groups. On one hand, the carbohydrates can modify intrinsic properties of a protein by altering its size, charge, solubility, accessibility, structure, or dynamic properties. On the other hand, the glycans themselves may be specifically recognized by carbohydrate-binding proteins and thus participate in adhesion processes and signal transduction.

2 Modulatory and Structural Roles of Glycans

In the following sections some modulatory and structural roles of glycans are discussed. However, it should be kept in mind that the chosen classification of the roles of glycans is somehow arbitrary since the individual effects often cannot be separated from each other. Thus, varied stability or biological function of a protein upon glycosylation is often a result of altered structural or dynamic parameters caused by the carbohydrates.

2.1 Modulation of Physicochemical Properties

The modification of physicochemical properties of proteins by attached carbohydrates is often observed, especially in glycoproteins with high carbohydrate content [1,7,12,26]. Sialylated or sulfated glycans, e.g., change the overall charge of a protein and increase its solubility. This is important for the highly glycosylated mucins (high sialic acid content) and the highly sulfated proteoglycans. Contained in mucous secretions of most epithelial cells, they can provide a gelation function due to their ability to retain water. Thus, they function as lubricants and protection for epithelial surfaces and mediate transport. Examples include synovia, buffering excessive gastric acid, and transport of chyme. In the extracellular matrix, proteoglycans provide elasticity and tensile strength. Furthermore, both mucins [27,28] and proteoglycans [29,30,31,32,33] act as adhesion molecules in numerous cell-cell, cell-matrix, and cell-microbe recognition events.

Antifreeze glycoproteins (AFGPs) [34,35,36,37,38] circulate in the blood of Antarctic fish and enable them to avoid freezing in their perpetually icy environment where the temperature is frequently as low as -1.9°C . These mucin-type *O*-glycoproteins are composed of repeats of the glycotriptide unit $(\text{Ala-Ala-}[\text{Gal}(\beta 1-3)\text{GalNAc}(\alpha 1-O)]\text{Thr})_n$ (• Fig. 1). Eight distinct fractions of these proteins have been isolated (AFGP 1–8) with the number n of glycotriptide

■ Figure 1

Structure of the glycotriptide repeating unit $(\text{Ala-Ala-}[\text{Gal}(\beta 1-3)\text{GalNAc}(\alpha 1-O)]\text{Thr})_n$ contained in antifreeze glycoproteins

repeats ranging from 50 (molecular mass 33 kDa) to 4 (2.6 kDa). In smaller-sized AFGPs, proline replaces some of the alanine residues following threonine. Antifreeze glycoproteins function in a noncolligative manner by binding to and inhibiting the growth of ice crystals that enter the fish and maybe also by preventing the nucleation of ice crystals. This results in a freezing point depression without an appreciable change in the melting point. The difference between the melting and freezing temperatures is termed thermal hysteresis and is used as a measure for the magnitude of the antifreeze activity. In addition to thermal hysteresis, AFGPs effect alteration of the morphology of ice crystals into a hexagonal bipyramid. Studies revealed that chemical modification (acetylation, periodate oxidation, and complexation with borate but not oxidation of the two primary hydroxyls with galactose oxidase to give the bisaldehyde) or removal of the sugar residues results in a complete loss of antifreeze activity [35,38]. Recently, Nishimura et al. presented a detailed study on synthetic AFGPs composed of one to seven glycosylated Ala-Thr-Ala repeating units [39]. It was shown that a minimum of two

■ Figure 2

a Superposition of the 25 lowest energy structures of $(\text{Ala}[\text{Gal}(\beta 1-3)\text{GalNAc}(\alpha 1-0)]\text{Thr-Ala})_3$ calculated from NMR-based constraints. The peptide backbone is navy, and the carbohydrate moieties are royal blue. b Structure closest to the average of the 25 best calculated models: yellow, methyl carbon; white, carbon in carbohydrate; gray, carbon in peptide side chains except for methyl carbon; blue, nitrogen; red, oxygen; the *N*- and *C*-terminal ends are identified. The peptide backbone folds into a left-handed helix similar to the polyproline type II helix. The disaccharides point to one side of the amphiphatic helix, thus constructing a hydrophilic face, while the Ala-CH₃ groups and acetyl methyl groups in the GalNAc residues are clustered into the hydrophobic face of the molecule. (Reprinted from [39] with permission from Wiley-VCH)

repeating glycopeptide units is necessary to obtain compounds that give rise to thermal hysteresis. All compounds including the monomeric glycotriptide, however, were capable to alter the morphology of ice crystals into hexagonal bipyramids. Varying the disaccharide part of oligomeric AFGP revealed three key motifs required for antifreeze activity, namely the *N*-acetyl group at the C-2 position of the reducing hexosamine, α -configuration of the *O*-glycosidic linkage between sugars and peptide chain, and the γ -methyl group of the threonyl residue. Strong evidence indicated that the peptide backbone conformation plays a critical role for antifreeze activity. Compounds lacking the mentioned key motifs either lost secondary structure or changed from a polyproline type II helix in the active compounds to a largely α -helical structure in the case of the glycosylated Ala-Ser-Ala repeating unit as determined by CD spectroscopy. Despite the absence of long- and medium-range NOE contacts, a surprisingly well-defined NMR-based structure could be calculated for the three repeating units containing glycopeptide (Ala-Ala-[Gal(β 1-3)GalNAc(α 1-O)]Thr)₃ (● Fig. 2). The authors propose that the amphiphatic nature of the observed left-handed helix is crucial for antifreeze activity. However, even though the study by Nishimura et al. presents a consistent picture, the situation for naturally occurring AFGPs seems to be more complicated and structural studies on these molecules have a long history of controversial interpretations [35].

Genetic studies have shown that the AFGPs present in the two geographically and phylogenetically distinct Antarctic notothenioids and Arctic cod have evolved independently, in a rare example of convergent molecular evolution [38]. Recently, a novel antifreeze glycoprotein has been isolated from *Pleuragramma antarcticum* which was shown to be a proteoglycan [39]. Beside these antifreeze glycoproteins, several non-glycosylated antifreeze proteins are known [34,35,36,37,38].

2.2 Protective and Stabilizing Functions

There is little doubt that the “coating” of oligosaccharides on many glycoproteins can protect the peptide backbone from recognition by proteases and antibodies [1,7,12,26]. An example is the decay accelerating factor (DAF, CD55) which is proteolyzed upon removal of its *O*-linked sugars [40]. Decoration of the surfaces of most types of cells with different kinds of glycoconjugates gives rise to the so-called glycocalyx which can present a substantial physical barrier. Glycosylation can also alter the heat stability of proteins which has been shown for two different β 1-3/4 glucanases from *Bacillus* species. Expression in *Saccharomyces cerevisiae* resulted in heavily glycosylated enzymes (carbohydrate content of about 45%) which were significantly more heat stable than their non-glycosylated counterparts expressed in *Escherichia coli* [41]. Such findings are of special interest for the industrial use of proteins. Other examples of the stabilizing function of carbohydrates [42,43,44] are ovine submaxillary mucin [45], an isolated tailpiece from human serum immunoglobulin M [46], RNase B [47,48,49], human CD2 [50], and the protease inhibitor PMP-C [51]. In these cases, the glycans had little overall effect on the conformation of the proteins but rather globally decreased the dynamic fluctuations of the glycoproteins, as revealed by NMR spectroscopy. The stabilizing function can be achieved even by a single carbohydrate unit [50,51].

2.3 Modulation of Biological Activity

There are many examples for the ability of carbohydrates to modulate the biological activity of functional proteins [1,7,12]. Bovine pancreas ribonuclease (RNase) for example occurs in unglycosylated (RNase A) and glycosylated (RNase B) forms, the latter being a set of nine different glycoforms of the high mannose type ($\text{Man}_5\text{--}_9[\text{GlcNAc}]_2$) with respect to the only *N*-glycosylation site (Asn-34) of the protein [7]. Using double stranded RNA as substrate, it was shown that RNase A is more than three times as active as RNase B [49]. Furthermore, enzyme activities of several glycoforms of RNase, prepared by exomannosidase treatment of naturally occurring RNase B, have been determined and may be ranked in terms of decreasing activity as: RNase A > RNase ($\text{Man}_0[\text{GlcNAc}]_2$) > RNase ($\text{Man}_1[\text{GlcNAc}]_2$) > RNase ($\text{Man}_5[\text{GlcNAc}]_2$) > RNase B. These differences in activities were attributed to an overall increase in dynamic stability of the protein upon glycosylation and to steric hindrance between the oligosaccharides and the RNA substrate which is also supported by molecular modeling [47,48,49,52]. The dynamic properties of the protein were studied by NMR spectroscopic determination of proton/deuterium exchange rates of the various glycoforms. The presence of the glycan reduces solvent access to many regions of the peptide backbone both close to and remote from the glycosylation site as much as 30 Å away (● Fig. 3), suggesting

■ Figure 3

The presence of the glycan (light blue) on ribonuclease B reduces the amide proton/deuterium exchange rates compared with ribonuclease A for extensive regions of the peptide backbone (shown in red) both local to and remote from the glycosylation site [47]. The glycoprotein structure is based on the crystal structure of ribonuclease B [53] and one of the structures of $\text{Man}_9\text{GlcNAc}_2$ determined by NMR and molecular dynamics. (Reprinted from [44] with permission from Elsevier), [54]

that it reduces fluctuations of the backbone. This was confirmed by circular dichroism (CD) studies that indicated that the glycan has a small stabilizing effect on the peptide fold.

Tissue plasminogen activator (tPA) is another enzyme whose activity is affected by glycosylation [7]. tPA is a serine protease which converts plasminogen into plasmin which itself is a serine protease with fibrinolytic activity. There are two major classes of glycoforms of naturally occurring tPA. Type I contains three *N*-linked glycans at Asn-117, Asn-184, and Asn-448, whereas type II has only two, at Asn-117 and Asn-448. Variable occupancy of Asn-184 affects the population of glycoforms at Asn-448. Plasminogen is also a mixture of two major glycoforms containing one *O*-glycan and one *N*-glycan (type 1) or one *O*-glycan (type 2). For an efficient proteolytic activity of tPA, formation of a ternary complex with plasminogen and fibrin is required. Rate of formation and turnover of the complex is dependent on the glycosylation site occupancy of both tPA and plasminogen [55]. Thus, clot lysis occurs 2–3 times faster with type II tPA in combination with type 2 plasminogen versus participation of type I tPA and type 1 plasminogen. Both the *N*-glycan linked to Asn-184 in type I tPA and the *N*-glycan contained in type 1 plasminogen reduce the enzymatic activity of the serine proteases. In contrast to this tuning function of the *N*-glycans, it has been shown that the *O*-glycan contained in plasminogen is crucial for its proteolytic activity.

Glycosylation of a ligand can also mediate such an on-off or switching effect. Deglycosylation of the hormone human β chorionic gonadotropin (β -HCG) for example leads to a species which is still able to bind to its receptor with similar affinity but fails to stimulate adenylate cyclase [1]. In most cases, however, such effects of glycosylation are incomplete, i. e., the carbohydrates provide a means of tuning the primary function of the proteins. From numerous examples studied, it was concluded that the relatively large *N*-linked glycans generally down-modulate the activities of enzymes and signal molecules whereas *O*-glycosylation can result in both a downregulation and an upregulation [10,56,57].

2.4 Influence on Peptide Secondary Structure

Glycosylation can affect protein structure in several capacities [7,9,10,42,43,44,58,59]. Structural roles of glycans associated with protein stability and the regulation of protein function have already been mentioned in the preceding sections. Oligosaccharides attached to matrix molecules like collagens and proteoglycans are important in the physical maintenance of tissue structure, porosity, and integrity. *N*-Glycosylation is a co-translational process that is believed to play a major role in the initiation of correct folding of the nascent polypeptide chain in the rough endoplasmic reticulum [9,43,58,60,61]. Indeed, numerous examples exist in which removal of certain *N*-glycosylation sites by site-directed mutagenesis results in improper folding of glycoproteins [9,58,62,63,64,65,66]. The impact of *N*-glycosylation on the conformation of model peptides was nicely demonstrated by Imperiali et al. in a series of publications using time-resolved fluorescence resonance energy transfer (FRET) [67] and NMR techniques [68,69]. The examined peptides were derived from the A282-A288 sequence (Ile-Thr-Pro-Asn-Gly-Ser-Ile) of the hemagglutinin glycoprotein from influenza virus containing the critical Asn-A285 glycosylation site. This sequence represents a β -turn surface loop in the native protein which is of considerable interest as this motif is a common feature among the *final* structures of glycosylation sites in many glycoproteins.

■ Figure 4

Peptide 1 and glycopeptides 2–6 synthesized in order to probe the influence of asparagine-linked glycosylation on peptide secondary structure [67,68,69]

In the FRET study [67], fluorescently labeled analogues of peptide 1 and glycopeptide 2 (● Fig. 4) with a dansyl group at N_8 of Orn were examined. The study revealed that glycosylation of 1 with a chitobiosyl moiety ($\rightarrow$ 2) promoted the adoption of a more compact peptide secondary structure. Subsequent 2D ^1H NMR investigations in aqueous solution [68] supported this analysis and indicated that peptide 1 adopts an open and extended Asx-turn conformation prior to glycosylation whereas glycopeptide 2 exhibits a compact type I β -turn conformation (● Fig. 5), quite similar to that observed in the final native protein structure. Both the FRET and NMR studies provide direct evidence that glycosylation of peptide 1 with chitobiose, a disaccharide representing the first two *N*-acetylglucosamine residues of the native tetradecasaccharide (cf. ● Chap. 8.1), induces a conformational switch in the peptide backbone. This observation is important with respect to the role glycosylation plays in the correct folding of glycoproteins. It was suggested that *N*-linked glycosylation may serve as a critical trigger to help the polypeptide chain to adopt a conformation that is populated in the native folded protein, but not in the nascent unmodified sequence.

Despite the strong influence of glycosylation on peptide conformation, no specific interactions between the chitobiose moiety and the peptide backbone were detected in the NMR analysis in aqueous solution. Therefore, it was proposed that the conformational change observed upon

■ Figure 5

Glycosylation-induced conformational switching from an (a) Asx-turn to a (b) type I β -turn [68]

glycosylation results from either a steric effect in which the disaccharide alters the conformational space available to the peptide or from a modulation of the local water structure that influences the environment that the peptide experiences [68]. To address the important question if chitobiose may be replaced with other disaccharides, glycopeptides **3–6**, containing the saccharides Glc(β 1–4)GlcNAc, GlcNAc, GlcNAc(β 1–4)Glc, and Glc(β 1–4)Glc, respectively, were examined (● Fig. 4) [69]. In all cases, less well-ordered peptide conformations as compared to **2** were determined by NMR analysis. The study revealed that the *N*-acetyl group of the proximal sugar is critical for maintaining a β -turn conformation. Surprisingly, the *N*-acetyl group of the distal sugar also plays an important role in rigidifying both the saccharide and the peptide.

Recently, Imperiali et al. reported a synergistic experimental and computational study of a glycopeptide analogous to glycopeptide **2** but with the chitobiose moiety α -*N*-glycosidically linked to asparagine [70]. Thereby, the effect of the stereochemistry of the carbohydrate-peptide linkage on glycopeptide structure could be evaluated. It was shown that only the β -linked glycopeptide **2** adopted a type I β -turn whereas the α -anomer adopted an Asx-turn-like conformation comparable to that of the unmodified peptide **1**. In this regard, it is worth mentioning that statistical analyses of the conformation of *N*-glycosylation sites revealed that only about a quarter of glycosylated asparagines are located at a β -turn [71,72,73]. This could mean that the induction of a β -turn structure by *N*-glycosylation is of transient character and of major importance only during the early stages of the folding process [39]. A lectin-mediated mechanism by which *N*-glycans participate in quality control of protein folding will be discussed in ● Sect. 3.3.2.

An influence of *N*-glycosylation on peptide conformation has also been observed by Danishefsky et al. upon attachment of the non-natural trisaccharide Gal(β 1–6)Gal(β 1–6)GlcNAc β -*N*-glycosidically to the Asn side chain of the model peptide H-Ala-Leu-Asn-Leu-Thr-OH [74]. Whereas the unglycosylated peptide failed to manifest any appreciable secondary structure, the glycopeptide was assumed to exist in an equilibrium between an ordered and a random state. When carried out at -12°C using a 90:10 mixture of H_2O /acetone- d_6 , NMR analysis

revealed nuclear Overhauser enhancement (NOE) effect crosspeaks between the methyl of the sugar *N*-acetyl group and backbone amide protons.

In contrast to *N*-glycosylation, *O*-linked glycosylation is an entirely post-translational and post-folding event in mammalian cells. Nevertheless, *O*-Glycans may have major impact on the secondary, tertiary, and quaternary structure of fully folded proteins [16]. Otvos et al. showed that glycosylation distorts the α -helicity of an epitope on the rabies virus glycoprotein, and that *O*-glycosylation of threonine is more effective in perturbing the secondary structure of the peptide than *N*-glycosylation of asparagine [75]. In several NMR studies it has been shown that the peptide backbone of model peptides responds to *O*-glycosylation as evidenced by changes in sequential amide-amide NOE interactions [76,77,78,79,80]. The examples include the mucin-type GalNAc(α 1-O)Thr [76,77,78] as well as the GlcNAc(β 1-O)Thr linkages [79,80]. In all cases, the NMR data, which were supplemented in part by CD and fluorescence measurements [78,79], molecular modeling calculations [78,79,80], and chemical evidence [80], were indicative for a glycosylation-induced conformational change from a random structure to a turn-like structure. The conformational response is further modulated by whether the sugar component is a mono-, di-, or oligosaccharide [77,80].

In recent times, more detailed conformational analyses of *O*-glycopeptides have been published by several groups [81,82,83,84,85,86,87,88]. Very recently, Griesinger and Kunz and coworkers reported the effect of *O*-glycosylation with the α (2,6)-sialyl-T antigen on the conformational propensities of a 20-residue peptide representing the full length tandem repeat sequence of the human mucin MUC1 [89]. The peptide contained both the GVTSP sequence, which is an effective substrate for GalNAc transferases, and the PDTRP fragment, a known epitope recognized by several anti-MUC1 monoclonal antibodies. NMR experiments were carried out in a H₂O/D₂O mixture (9 : 1) close to physiological conditions (25 °C, pH 6.5). Cluster analysis of the conformational ensemble yielded a rod-like structure with the GalNAc methyl group oriented towards the C-terminus, similar to results obtained by Danishefsky et al. [81].

3 Involvement of Glycans in Recognition Events

Oligosaccharides have an enormous information-storing potential, being substantially higher than that of oligopeptides and oligonucleotides [90]. Thus, it is not surprising that the carbohydrates contained in glycoconjugates, beside their role in modulating intrinsic properties of glycoproteins, can act as recognition markers in numerous physiological and pathological processes [1,7,12,15,91,92,93,94,95,96,97,98,99,100]. Responsible for deciphering the encoded information are carbohydrate-binding proteins. They can be subdivided into antibodies, enzymes involved in sugar utilization and glycoconjugate turnover, and carbohydrate-binding proteins which are neither immunoglobulins nor enzymes. The latter have been referred to as lectins [101,102,103,104,105,106,107]. Carbohydrate-lectin interactions play a crucial role in many cellular recognition processes including clearance of glycoproteins from the circulatory system, control of intracellular traffic of glycoproteins, bacterial and viral adhesion to host cells, recruitment of leukocytes to inflammatory sites, cell interactions in the immune system, and tumor metastasis just to name a few. In the following sections some of these processes will be highlighted.

3.1 Carbohydrate Recognition by Blood Group Antibodies

The discovery of the ABO blood group system 100 years ago by Landsteiner et al. [108] was based on the observation that humans could be divided into different groups according to the presence or absence of serum constituents that would agglutinate red cells isolated from other humans. Although they were not aware of the underlying glycan basis, the work of Landsteiner and colleagues laid the basis for the safe transfusion of blood from one individual to another [109]. Today we know that the agglutinating serum constituents are antibodies and that their cognate antigens are oligosaccharides whose structures are genetically polymorphic [12,110,111,112,113]. The first discovery of these antigens on the surface of human erythrocytes led them to be classified as “blood-group antigens”. However, they are also found in human secretions and mucosal tissues and it was suggested that they are more accurately defined as “histo-blood group antigens” [114].

The A, B, and H antigens (the latter being the one expressed by blood group O individuals) are formed by sequential action of distinct glycosyltransferases on different precursor saccharides, most often terminal residues of glycolipids and proteins. According to the nature of these peripheral disaccharide core structures on which the blood group antigens are synthesized, different types are distinguished. In type-1 structures, the antigen synthesis starts from $\text{Gal}(\beta 1-3)\text{GlcNAc}(\beta)$. Similarly, modification of $\text{Gal}(\beta 1-4)\text{GlcNAc}(\beta)$, $\text{Gal}(\beta 1-3)\text{GalNAc}(\alpha)$, and $\text{Gal}(\beta 1-3)\text{GalNAc}(\beta)$ leads to type-2, type-3, and type-4 antigens, respectively. Fucosylation of these precursors by an $\alpha 1-2$ fucosyltransferase encoded by the *H* or the *Secretor* locus produces the blood group H determinant represented by the disaccharide $\text{Fuc}(\alpha 1-2)\text{Gal}(\beta)$ (Scheme 1). Further glycosylation by an $\alpha 1-3$ *N*-acetylgalactosaminyltransferase ($\alpha 1,3$ -GalNAcT) (corresponding to the A allele of the *ABO* locus) or an $\alpha 1-3$ galactosyltransferase ($\alpha 1,3$ -GalT) (corresponding to the B allele of the *ABO* locus) then leads to $\text{GalNAc}(\alpha 1-3)[\text{Fuc}(\alpha 1-2)]\text{Gal}(\beta)$ (blood group A determinant) and $\text{Gal}(\alpha 1-3)[\text{Fuc}(\alpha 1-2)]\text{Gal}(\beta)$ (blood

■ Scheme 1

Biosynthesis of the blood group antigens A and B (minimal determinant structures)

Table 1

The ABH(O) system of antigens on erythrocytes and antibodies and glycosyltransferases in plasma

Geno- type	Phenotype	Blood group antigens on red cells (minimal determinant structure)	Antibodies in plasma	Glycosyltransferases in plasma
AA AO	A	A	anti-B	α 1,3-GalNAcT
BB BO	B	B	anti-A	α 1,3-GalT
AB	AB	A and B	–	α 1,3-GalNAcT, α 1,3-GalT
OO	O	H	anti-A, anti-B	–

Table 2

Selected antigenic determinants (LacNAc = Gal(β 1–4)GlcNAc)

Determinant structure	Name
Gal(β 1–3)[Fuc(α 1–4)]GlcNAc	Blood group Lewis a (Le ^a)
Neu5Ac(α 2–3)Gal(β 1–3)[Fuc(α 1–4)]GlcNAc	Sialyl Lewis a (sLe ^a)
Fuc(α 1–2)Gal(β 1–3)[Fuc(α 1–4)]GlcNAc	Blood group Lewis b (Le ^b)
Gal(β 1–4)[Fuc(α 1–3)]GlcNAc	Lewis x (Le ^x)
Neu5Ac(α 2–3)Gal(β 1–4)[Fuc(α 1–3)]GlcNAc	Sialyl Lewis x (sLe ^x)
Fuc(α 1–2)Gal(β 1–4)[Fuc(α 1–3)]GlcNAc	Lewis y (Le ^y)
[LacNAc(β 1–3)] _n	Blood group i
LacNAc(β 1–3)[LacNAc(β 1–6)]LacNAc	Blood group I
GalNAc(α 1–O)Ser/Thr	Tn
Gal(β 1–3)GalNAc(α 1–O)Ser/Thr	T
Fuc(α 1–2)Gal(β 1–3)GalNAc(β 1–3)Gal(α 1–4)Gal(β 1–4)Glc	Globo H

group B determinant), respectively. The ABO classification is based on the presence or absence of the A and B antigens and the two antibodies anti-A and anti-B which always occur in the plasma when the corresponding antigen is missing (Table 1). With rare exceptions the H antigen is expressed on the cells of all blood group O individuals, but in persons belonging to phenotypes A, B, and AB there is complete or partial masking of H activity. Despite the vast accumulated knowledge of serology, chemistry, and genetics of the blood group structures, it is still not possible to assign a clearly defined physiological function to the ABO locus.

Beside the ABO blood group system a number of additional carbohydrate-based antigenic determinants are known (Table 2), which are, however, not only recognized by antibodies. Some of them are more highly expressed on tumor cells and therefore can be described as tumor-associated antigens [115]. Synthetic tumor-associated antigens such as Lewis y or Globo H have been used in the development of antitumor vaccines [116,117].

3.2 Carbohydrate-Modifying Enzymes

Of the many enzymes involved in carbohydrate metabolism, glycosyl transferases and glycosidases are of special interest for the biosynthesis and processing of glycoconjugates. The

■ Scheme 2

Enzymatic galactosyl transfer from UDP-Gal to the 4-position of GlcNAc catalyzed by β 1-4 galactosyltransferase (β 1,4-GalT)

glycosyl transferases of the Leloir pathway [118] which are responsible for the synthesis of most glycoproteins and glycolipids in mammalian systems [119] are membrane-bound or membrane-associated enzymes. They utilize as glycosyl donors monosaccharides which are activated as glycosyl esters of nucleoside mono- and diphosphates (so-called sugar nucleotides) to transfer a glycosyl moiety to an acceptor [120,121,122]. The eight sugar nucleotides applied for the synthesis of most oligosaccharides are uridine 5'-diphosphoglucose (UDP-Glc), UDP-*N*-acetylglucosamine (UDP-GlcNAc), UDP-galactose (UDP-Gal), UDP-*N*-acetyl-galactosamine (UDP-GalNAc), UDP-glucuronic acid (UDP-GlcA), guanosine 5'-diphosphomannose (GDP-Man), GDP-fucose (GDP-Fuc), and cytidine 5'-monophospho *N*-acetylneuraminic acid (CMP-Neu5Ac). Non-Leloir transferases typically utilize glycosyl phosphates as donors. For each sugar nucleotide glycosyl donor, several Leloir transferases exist each of which acts on a different acceptor in a regio- and stereospecific manner. The transferases are generally considered to be specific for a given glycoside bond, however, some deviations from this picture of absolute specificity have been observed, both in the glycosyl donors and acceptors. β 1-4 Galactosyltransferase for example catalyzes the transfer of galactose from UDP-Gal to the 4-position of β -linked GlcNAc residues to produce the Gal(β 1-4)GlcNAc substructure (● Scheme 2). In the presence of lactalbumin, glucose is the preferred acceptor, resulting in the formation of lactose, Gal(β 1-4)Glc [123].

Many glycosyltransferases have been employed in the *in vitro* synthesis of oligosaccharides [124,125]. Since the nucleoside mono- or diphosphates released during the reaction, e. g. UDP, usually cause product inhibition, *in situ* regeneration cycles have been developed in which they are transformed back to the sugar nucleotides [126]. Alternatively, the nucleoside mono- or diphosphates may be degraded by use of alkaline phosphatase [127].

Glycosidases (glycosyl hydrolases, EC 3.2.1) [128] catalyze the hydrolysis of glycosidic bonds. For the naturally occurring polysaccharides cellulose and starch typical rate constants up to 1000 s^{-1} are observed. Taking the half-lives for spontaneous hydrolysis of these materials (about 5 million years [129]) into account, these enzymes achieve rate enhancements of up to 10^{17} -fold, placing them amongst the most proficient of enzymes. In nature, glycosidases play important roles ranging from the degradation of polysaccharides as food sources

■ Scheme 3

Mechanisms of (a) inverting and (b) retaining glycosidases. Both mechanisms involve transition states with substantial oxocarbenium character present at the anomeric center

through to the trimming process of glycoproteins. The hydrolysis reaction can occur with either inversion or retention of anomeric configuration. Depending on the stereochemical outcome, two different mechanisms have been suggested 1953 by Koshland [131] and later on refined [131,132,133,134]. Both mechanisms involve a pair of carboxylic acids at the active site (● Scheme 3). In inverting glycosidases, these residues are approximately 10 Å apart from each other on average. Reaction occurs by a single-displacement mechanism with one carboxylic acid acting as a general base activating a water molecule and the other as a general acid (● Scheme 3a). In retaining enzymes, the carboxyl groups are only approximately 5.5 Å apart and the hydrolysis proceeds by a double-displacement mechanism involving a covalent glycosyl-enzyme intermediate (● Scheme 3b). In both mechanisms, substantial oxocarbenium character has been shown to be present at the anomeric center in each transition state. The design of transition-state analogs mimicking the shape and charge of the oxocarbenium species has been extensively used to arrive at glycosidase inhibitors [135]. An example of a designed inhibitor which was approved by the Food and Drug Administration (FDA) in 2000 for the treatment of influenza virus A and B infections is the sialidase inhibitor zanamivir (RelenzaTM) [136]. The search for a drug with improved pharmacokinetic properties led to the zanamivir mimic oseltamivir that came on the market in 2000 under the name of TamifluTM which is the phosphate salt of an ethyl ester prodrug [137].

In contrast to glycosyltransferases, glycosidases show broader, sometimes overlapping substrate specificities, especially with respect to the aglycon. Endoglucanases, for example, whilst typically considered to be cellulases, are also active to various degrees on xylan, xyloglucan, β -glucan, and various artificial substrates. Retaining glycosidases, run in the transglycosyla-

tion mode, have also been used for the construction of glycosidic bonds, albeit in moderate yields, the major problem being the thermodynamically favored hydrolysis of the reaction product [125,138]. Withers et al. could overcome this problem with a mutated β -glucosidase, wherein the nucleophilic carboxyl had been substituted with alanine [139]. The mutated glycosidase (a so-called glycosynthase [140,141,142]) folds correctly but lacks hydrolase activity since it cannot form the requisite α -glycosyl-enzyme intermediate. It possesses, however, high transglycosylation activity. Employing an activated α -glycosyl fluoride having the opposite anomeric configuration to that of the normal substrate, the ligation to a suitable acceptor sugar bound in the aglycon pocket is catalyzed. Once the oligosaccharide is formed, it cannot be hydrolyzed by the enzyme, thus allowing yields of over 90% to be achieved.

3.3 Carbohydrate-Lectin Interactions

3.3.1 Classification of Lectins

Lectins [102,103,104,105,143,144,145,146] are carbohydrate-binding proteins other than immunoglobulins without enzymatic activity towards the recognized sugars [101]. They are present throughout nature including the animal kingdom and the microbial world. Originally, they were classified according to their monosaccharide binding specificity. However, with the advent of molecular cloning a more consistent classification based on amino acid sequence homology and evolutionary relatedness emerged [147]. Whereas the biological function of plant lectins is still unclear, their use as biochemical tools has made an enormous contribution to the understanding of the structures and functions of carbohydrate structures in animal cells. Today, we have a much more detailed picture of the function of animal lectins than of plant lectins. Table 3 shows selected examples of the 13 currently known families of animal lectins [12,102,105].

C-Type lectins were named after their requirement of calcium ions for recognition. They are all characterized by an extracellular carbohydrate recognition domain (CRD) and bind a diver-

Table 3

Selected families of animal lectins classified according to known sequence homologies^a [12,102,105]

Lectin family	Defining structural motif	Carbohydrate ligand	Calcium dependence
C-type (includes selectins, collectins, endocytic lectins)	Conserved CRD	Variable	Yes (most)
Galectins (formerly S-type)	Conserved CRD	β -galactosides	No
P-type	Unique repeating motif	Mannose-6-phosphate on high-mannose-type N-glycans	Variable
I-type (includes Siglec family)	Immunoglobulin-like CRD	Variable (Siglecs: sialic acids)	No
Calnexin, calreticulin, calmeglin	Homology with each other	Glucosylated high-mannose-type N-glycans in the ER	Yes
Hyaluronan-binding proteins	Homologue CRD	Hyaluronan chains	No

^a Abbreviations: CRD, carbohydrate recognition domain; ER, endoplasmic reticulum

sity of sugars. However, not all calcium-requiring lectins are C-type lectins, as exemplified by calnexin and calreticulin which recognize glucose residues on newly synthesized *N*-glycoproteins. Galectins, formerly named S-type lectins due to their dependency on free thiols for full activity, are soluble β -galactoside-specific lectins that combine preferentially with lactose and *N*-acetylglucosamine. P-Type lectins bind mannose-6-phosphate as their ligand and I-type lectins share a common immunoglobulin-like CRD. Another class of evolutionarily very ancient circulating soluble lectins are the pentraxins. They are characterized not so much by primary sequence homologies, but by a pentameric arrangement of their subunits and a probable role in the primary host immune response.

Multivalency appears to play an important role in lectin-mediated interactions [106,148,149,150,151], and many lectins are found to recognize individual carbohydrate epitopes only with low affinity. Indeed, most lectins are either intrinsically multivalent because of their defined multisubunit structure or by virtue of having multiple CRDs within a single polypeptide or they can become functionally multimeric by noncovalent association or by clustering on cell surfaces. High-avidity binding can result from multiple interactions of adequately presented low-affinity single sites, and this appears to be a common mechanism of modulating lectin function in vivo [144,152,153,154].

The following sections will focus on selected, well-characterized carbohydrate-lectin recognition processes.

3.3.2 Lectin Control of Protein Folding

N-Linked glycan moieties of glycoproteins which are incorporated co-translationally in the endoplasmic reticulum (ER) during biosynthesis are in many cases essential for proper folding of the protein. Their conformational influence on nascent or newly synthesized glycoproteins may be either direct by inducing local structure elements, such as β -turns, (cf. **● Sect. 2.4**) or indirect by interaction with calnexin or calreticulin within the so-called calnexin-calreticulin-cycle (**● Fig. 6**) [92,96,155,156]. Membrane-bound calnexin (CNX) and soluble calreticulin (CRT) are homologous lectins found in the ER of nearly all eukaryotes. They bind to a monoglucosylated form of *N*-glycoproteins after two of the glucoses have been trimmed away by glucosidases I and II. Binding by calnexin or calreticulin retains the unfolded glycoprotein in the ER, permitting foldases to help folding into the correct 3D arrangement. Folding factor ERp57 is a thiol oxidoreductase homologue of protein disulfide isomerase that also binds to both calnexin and calreticulin and thereby is exposed to the substrate glycoprotein. If the glycoprotein contains cysteines, the formation of proper disulfide bonds is catalyzed through the formation of transient mixed disulfides with ERp57. Association of the glycoprotein with calnexin and calreticulin is dynamic and exposes the single terminal glucose residue to hydrolysis by glucosidase II. If the deglycosylated glycoprotein is incorrectly or incompletely folded, it is either reglucosylated by UDP-glucose glycoprotein glucosyltransferase, which has the remarkable function of a folding sensor, or degraded. Reglucosylation allows another cycle of binding by calnexin or calreticulin and interaction with ERp57. Once correctly folded, the protein is no longer recognized by the glucosyltransferase and, therefore, does not bind back any more to calnexin and/or calreticulin but can leave the ER. Exit of certain glycoproteins from the ER to the Golgi complex is assisted by another membrane-bound lectin, ERGIC-53, which binds to mannose residues.

■ Figure 6

The calnexin-calreticulin-cycle (cf. text). Triangles represent glucose residues (Reprinted from [92] with permission from AAAS)

The calnexin-calreticulin-cycle promotes correct folding, inhibits aggregation of folding intermediates, and provides quality control by preventing incompletely folded glycoproteins from exiting to the Golgi complex. It seems to be essential *in vivo*. Transgenic mice devoid of calreticulin die on embryonic day 18 [157].

3.3.3 Clearance and Targeting of Glycoproteins

Some effects of glycosylation on the stability of proteins have already been mentioned in Sect. 2.2 and they can presumably affect their half-life in single cells. In the intact organism, recognition of glycan structures by certain receptors can result in removal of the glycoconjugate or even a whole cell from the circulation. This has first been observed in the late 1960s by Ashwell and coworkers who serendipitously found that desialylation of glycoproteins resulted in significantly shorter serum half-lives [158,159]. This led to the characterization of the asialoglycoprotein receptor [160] (a hepatic lectin of the C-type) which has recently been crystallized [161] and not only recognizes desialylated, i.e. galactose terminated, *N*-linked oligosaccharides but also GalNAc structures as they occur on some *O*-linked glycans. These findings were traditionally interpreted as representing a physiological clearing mechanism for glycoproteins [160] which has, however, not yet proven beyond doubt [162]. Another exam-

ple of a receptor responsible for glycoprotein clearing is the GalNAc-4-sulfate receptor. Rapid removal of luteinizing hormone, for example, which contains this (*N*-linked) sugar is important to generate a circadian rhythm [163].

The mannose 6-phosphate (Man6P) receptors are the best understood examples of receptors responsible for intracellular trafficking of glycoproteins [164,165,166]. These P-type lectins mediate the routing of lysosomal hydrolases from the trans-Golgi network to their final destination in the lysosomes. Two such receptors are known, one cation-dependent and of low molecular weight (ca. 45 kDa) which has been crystallized [165], the other cation-independent and of high molecular weight (ca. 275 kDa). The targeting is mediated by recognition of Man6P residues on oligomannose-type *N*-glycans of lysosomal enzymes by the Man6P receptors. A defect in the synthesis of the Man6P markers, caused by a deficiency of GlcNAc-phosphotransferase (the first enzyme in the mannose phosphorylation pathway), results in I-cell disease (also called mucopolidosis II or MLII), an inherited lysosomal storage disease.

3.3.4 Leukocyte Trafficking

The leukocyte trafficking to inflammatory sites is a highly regulated multistep process, referred to as the inflammatory cascade [167]. The initial events, the tethering and rolling of leukocytes along the vascular endothelium, are mediated by the interaction of a family of adhesion molecules, termed selectins, and their carbohydrate containing ligands [168,169,170]. The cascade begins with the release of cytokines and other signaling molecules at the site of injury that stimulate the transient expression of E- and P-selectin on the endothelium surface. These C-type lectins bind to their ligands displayed on the circulating leukocytes and promote leukocyte adhesion to the stimulated endothelial cells. L-Selectin is constitutively expressed on leukocytes, and it recognizes its ligands on endothelial cells. The rolling leads to activation of integrins on the leukocytes that interact with their counter-receptors on endothelial cells (e.g. intercellular adhesion molecule-1, ICAM-1) and promote firm adhesion. This stronger interaction then allows emigration or extravasation of the leukocytes into the underlying damaged tissue. However, if too many leukocytes are recruited to the site of injury, normal cells can also be destroyed. This occurs in the condition known as septic shock, in chronic inflammatory diseases such as psoriasis and rheumatoid arthritis, and in the reperfusion tissue injury that occurs following a heart attack, stroke or organ transplant.

The selectins are membrane-bound proteins comprising five domains: a cytosolic tail that may play a role in signal transduction, a transmembrane domain, several complement regulatory domain repeats, an epidermal growth factor (EGF) domain, and an *N*-terminal, calcium-dependent C-type carbohydrate recognition domain (CRD). Both the EGF domain and the CRD are required for ligand binding, although the site of binding has been localized to the CRD. The tetrasaccharide sialyl Lewis x (sLe^x) Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc (Fig. 7) is the minimum structure recognized by all three selectins. It is bound, however, only weakly with *K_D* values of 0.7, 3.9, and 7.8 mM for binding to E-, L-, and P-selectin, respectively [171]. Since the interaction of the selectins with their natural glycoprotein ligands which contain sLe^x and modification thereof as terminating structures is much stronger, additional receptor-ligand contacts seem to be important for high-affinity binding. The natural ligand for L-selectin (GlyCAM-1) for example contains sLe^x sulfated at Gal-6, GlcNAc-6, or both positions and in the natural ligand for P-selectin, PSGL-1, the 19 amino acid *N*-terminus of the

■ Figure 7

The tetrasaccharide Neu5Ac(α 2-3)Gal(β 1-4)[Fuc(α 1-3)]GlcNAc (sialyl Lewis x), the minimum structure recognized the selectins. Indicated are functional groups involved in specific contacts to E-selectin (gray boxes) and P-selectin (light boxes) according to crystal structures [172]

protein which contains several sulfated tyrosine residues beside the *O*-linked, sLe^x-bearing glycan was found to be critical for binding. Recently, the crystal structures of human E- and P-selectin constructs containing the lectin and EGF (LE) domains co-complexed with sLe^x as well as the crystal structure of P-selectin LE in complex with the *N*-terminal domain of human PSGL-1 modified by both tyrosine sulfation and sLe^x have been reported [173]. This study offers a structural basis for affinity differences between sLe^x and E- and P-selectin, respectively and the high-affinity interaction between P-selectin and PSGL-1. In addition, it may well be that multivalent interactions between the selectins clustered on cell surfaces with multiple sLe^x residues presented on the highly glycosylated mucin-type counter-receptors contribute to high-affinity binding in vivo.

Inhibition of the selectin-ligand interactions is an attractive strategy for treating inflammation-related diseases [173,174,175]. As such, sLe^x has been intensively used as lead structure for development of anti-inflammatory drugs both in industry and academia [176,177,178].

References

- Varki A (1993) *Glycobiology* 3:97
- Lis H, Sharon N (1993) *Eur J Biochem* 218:1
- Fukuda M, Hindsgaul O (eds) (1994) *Molecular Glycobiology*. IRL Press, Oxford
- Karlsson K-A (1995) *Curr Opin Struct Biol* 5:622
- Montreuil J, Vliegthart JFG, Schachter H (eds) (1995) *Glycoproteins*. Elsevier, Amsterdam
- Montreuil J, Vliegthart JFG, Schachter H (eds) (1997) *Glycoproteins II*. Elsevier, Amsterdam
- Dwek RA (1996) *Chem Rev* 96:683
- Gabius H-J, Gabius S (eds) (1997) *Glycosciences: Status and Perspectives*. Chapman & Hall, Weinheim
- Rudd PM, Dwek RA (1997) *Crit Rev Biochem Mol Biol* 32:1
- Van den Steen P, Rudd PM, Dwek RA, Opdenakker G (1998) *Crit Rev Biochem Mol Biol* 33:151
- Drickamer K, Taylor ME (1998) *Trends Biochem Sci* 23:321
- Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds) (1999) *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Haines N, Irvine KD (2003) *Nat Rev Mol Cell Biol* 4:786
- Paulson JC, Blixt O, Collins BE (2006) *Nat Chem Biol* 2:238
- Lehle L, Strahl S, Tanner W (2006) *Angew Chem Int Ed* 45:6802

16. Taylor ME, Drickamer K (2006) *Introduction to Glycobiology*, 2nd ed. Oxford University Press, Oxford
17. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA (2007) *Annu Rev Immunol* 25:21
18. Seifert GJ, Roberts K (2007) *Annu Rev Plant Biol* 58:137
19. Wittmann V (ed) (2007) *Glycopeptides and Glycoproteins: Synthesis, Structure, and Application (Topics in Current Chemistry, vol. 267)*. Springer-Verlag, Berlin Heidelberg New York
20. Dube DH, Bertozzi CR (2003) *Curr Opin Chem Biol* 7:616
21. Prescher JA, Bertozzi CR (2005) *Nat Chem Biol* 1:13
22. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calareso D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong C-H, Paulson JC (2004) *Proc Natl Acad Sci USA* 101:17033
23. Feizi T, Chai W (2004) *Nat Rev Mol Cell Biol* 5:582
24. Shin I, Park S, Lee M-r (2005) *Chem Eur J* 11:2894
25. de Paz JL, Seeberger PH (2006) *QSAR Comb Sci* 25:1027
26. Sharon N, Lis H (1997) *Glycoproteins: Structure and Function*. In: Gabius H-J, Gabius S (eds) *Glycosciences: Status and Perspectives*. Chapman & Hall, Weinheim, p 133
27. Hanisch F-G (2001) *Biol Chem* 382:143
28. Hång HC, Bertozzi CR (2005) *Bioorg Med Chem* 13:5021
29. Mulloy B, Rider CC (2006) *Biochem Soc Trans* 34:409
30. Bülow HE, Hobert O (2006) *Annu Rev Cell Dev Biol* 22:375
31. Taylor KR, Gallo RL (2006) *FASEB J* 20:9
32. Imberty A, Lortat-Jacob H, Perez S (2007) *Carbohydr Res* 342:430
33. Bishop JR, Schuksz M, Esko JD (2007) *Nature* 446:1030
34. Yeh Y, Feeney RE (1996) *Chem Rev* 96:601
35. Wöhrmann APA (1998) *Soc Exp Biol Semin Scr* 66:266
36. Ewart KV, Lin Q, Hew CL (1999) *Cell Mol Life Sci* 55:271
37. Fletcher GL, Hew CL, Davies PL (2001) *Annu Rev Physiol* 63:359
38. Harding MM, Anderberg PI, Haymet ADJ (2003) *Eur J Biochem* 270:1381
39. Tachibana Y, Fletcher GL, Fujitani N, Tsuda S, Monde K, Nishimura S-I (2004) *Angew Chem Int Ed* 43:856
40. Coyne KE, Hall SE, Thompson S, Arce MA, Kinoshita T, Fujita T, Anstee DJ, Rosse W, Lublin DM (1992) *J Immunol* 149:2906
41. Olsen O, Thomsen KK (1991) *J Gen Microbiol* 137:579
42. Wyss DF, Wagner G (1996) *Curr Opin Biotechnol* 7:409
43. Imperiali B, O'Connor SE (1999) *Curr Opin Chem Biol* 3:643
44. Wormald MR, Dwek RA (1999) *Structure* 7:R155
45. Gerken TA, Butenhof KJ, Shogren R (1989) *Biochemistry* 28:5536
46. Wormald MR, Wooten EW, Bazzo R, Edge CJ, Feinstein A, Rademacher TW, Dwek RA (1991) *Eur J Biochem* 198:131
47. Joao HC, Scragg IG, Dwek RA (1992) *FEBS Lett* 307:343
48. Joao HC, Dwek RA (1993) *Eur J Biochem* 218:239
49. Rudd PM, Joao HC, Coghill E, Fiten P, Saunders MR, Opdenakker G, Dwek RA (1994) *Biochemistry* 33:17
50. Wyss DF, Choi JS, Li J, Knoppers MH, Willis KJ, Arulanandam ARN, Smolyar A, Reinherz EL, Wagner G (1995) *Science* 269:1273
51. Mer G, Hietter H, Lefèvre J-F (1996) *Nat Struct Biol* 3:45
52. Woods RJ, Edge CJ, Dwek RA (1994) *Nat Struct Biol* 1:499
53. Williams RL, Greene SM, McPherson A (1987) *J Biol Chem* 262:16020
54. Woods RJ, Pathiaseril A, Wormald MR, Edge CJ, Dwek RA (1998) *Eur J Biochem* 258:372
55. Mori K, Dwek RA, Downing AK, Opdenakker G, Rudd PM (1995) *J Biol Chem* 270:3261
56. Opdenakker G, Rudd PM, Ponting CP, Dwek RA (1993) *FASEB J* 7:1330
57. Opdenakker G, Rudd PM, Wormald M, Dwek RA, Van Damme J (1995) *FASEB J* 9:453
58. O'Connor SE, Imperiali B (1996) *Chem Biol* 3:803
59. Meyer B, Möller H (2007) *Top Curr Chem* 267:187
60. Imperiali B (1997) *Acc Chem Res* 30:452
61. Imperiali B, O'Connor SE, Hendrickson T, Kellenberger C (1999) *Pure Appl Chem* 71:777
62. Sodora DL, Cohen GH, Muggeridge MI, Eisenberg RJ (1991) *J Virol* 65:4424

63. Kern G, Kern D, Jaenicke R, Seckler R (1993) *Protein Sci* 2:1862
64. Duranti M, Gius C, Sessa F, Vecchio G (1995) *Eur J Biochem* 230:886
65. Feng W, Matzuk MM, Mountjoy K, Bedows E, Ruddon RW, Boime I (1995) *J Biol Chem* 270:11851
66. Lehmann S, Harris DA (1997) *J Biol Chem* 272:21479
67. Imperiali B, Rickert KW (1995) *Proc Natl Acad Sci USA* 92:97
68. O'Connor SE, Imperiali B (1997) *J Am Chem Soc* 119:2295
69. O'Connor SE, Imperiali B (1998) *Chem Biol* 5:427
70. Bosques CJ, Tschampel SM, Woods RJ, Imperiali B (2004) *J Am Chem Soc* 126:8421
71. Imberty A, Pérez S (1995) *Protein Eng* 8: 699
72. Petrescu A-J, Milac A-L, Petrescu SM, Dwek RA, Wormald MR (2004) *Glycobiology* 14: 103
73. Petrescu A-J, Wormald MR, Dwek RA (2006) *Curr Opin Struct Biol* 16:600
74. Live DH, Kumar RA, Beebe X, Danishefsky SJ (1996) *Proc Natl Acad Sci USA* 93:12759
75. Otvos L, Jr., Krivulka GR, Urge L, Szendrei GI, Nagy L, Xiang ZQ, Ertl HCJ (1995) *Biochim Biophys Acta* 1267:55
76. Andreotti AH, Kahne D (1993) *J Am Chem Soc* 115:3352
77. Liang R, Andreotti AH, Kahne D (1995) *J Am Chem Soc* 117:10395
78. Huang X, Smith MC, Berzofsky JA, Barchi JJ Jr (1996) *FEBS Lett* 393:280
79. Simanek EE, Huang DH, Pasternack L, Machajewski TD, Seitz O, Millar DS, Dyson HJ, Wong CH (1998) *J Am Chem Soc* 120:11567
80. Wu WG, Pasternack L, Huang DH, Koeller KM, Lin CC, Seitz O, Wong CH (1999) *J Am Chem Soc* 121:2409
81. Coltart DM, Royyuru AK, Williams LJ, Glunz PW, Sames D, Kuduk SD, Schwarz JB, Chen X-T, Danishefsky SJ, Live DH (2002) *J Am Chem Soc* 124:9833
82. Schuster O, Klich G, Sinnwell V, Kranz H, Paulsen H, Meyer B (1999) *J Biomol NMR* 14:33
83. Naganagowda GA, Gururaja TL, Satyanarayana J, Levine MJ (1999) *J Pept Res* 54:290
84. Kirnarsky L, Prakash O, Vogen SM, Nomoto M, Hollingsworth MA, Sherman S (2000) *Biochemistry* 39:12076
85. Grinstead JS, Koganty RR, Krantz MJ, Longenecker BM, Campbell AP (2002) *Biochemistry* 41:9946
86. Kirnarsky L, Suryanarayanan G, Prakash O, Paulsen H, Clausen H, Hanisch FG, Hollingsworth MA, Sherman S (2003) *Glycobiology* 13:929
87. von Mensdorff-Pouilly S, Kirnarsky L, Engelmann K, Baldus SE, Verheijen RH, Hollingsworth MA, Pisarev V, Sherman S, Hanisch FG (2005) *Glycobiology* 15:735
88. Möller H, Serttas N, Paulsen H, Burchell JM, Taylor-Papadimitriou J, Meyer B (2002) *Eur J Biochem* 269:1444
89. Dziadek S, Griesinger C, Kunz H, Reinscheid UM (2006) *Chem Eur J* 12:4981
90. Laine RA (1997) *The Information-Storing Potential of the Sugar Code*. In: Gabius H-J, Gabius S (eds) *Glycosciences: Status and Perspectives*. Chapman & Hall, Weinheim, p 1
91. Lee YC, Lee RT (1995) *Acc Chem Res* 28:321
92. Helenius A, Aeby M (2001) *Science* 291:2364
93. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA (2001) *Science* 291:2370
94. Lowe JB, Marth JD (2003) *Annu Rev Biochem* 72:643
95. Gabius H-J, Siebert H-C, André S, Jiménez-Barbero J, Rüdiger H (2004) *ChemBioChem* 5:740
96. Helenius A, Aeby M (2004) *Ann Rev Biochem* 73:1019
97. Freeze HH, Aeby M (2005) *Curr Opin Struct Biol* 15:490
98. Dube DH, Bertozzi CR (2005) *Nat Rev Drug Discovery* 4:477
99. Freeze HH (2006) *Nat Rev Genetics* 7:537
100. Haltiwanger RS (2002) *Curr Opin Struct Biol* 12:593
101. Barondes SH (1988) *Trends Biochem Sci* 13:480
102. Gabius H-J (1997) *Eur J Biochem* 243:543
103. Lis H, Sharon N (1998) *Chem Rev* 98:637
104. Kilpatrick DC (2000) *Handbook of Animal Lectins: Properties and Biomedical Applications*. Wiley, Chichester
105. Gabius H-J (ed) (2002) *Animal Lectins*, *Biochim Biophys Acta* 1572:163
106. Dam TK, Brewer CF (2002) *Chem Rev* 102:387
107. Sharon N (2007) *J Biol Chem* 282:2753
108. Landsteiner K (1900) *Zentralbl Bakt Parasitenkd Infektionskrankh Abt I* 27:357
109. Klein HG, Anstee DJ (2005) *Mollison's Blood Transfusion in Clinical Medicine*, 11th edn. Blackwell, Malden

110. Watkins WM (1995) Molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems. In: Montreuil J, Vliegthart JFG, Schachter H (eds) *Glycoproteins*. Elsevier, Amsterdam, p 313
111. Daniels GL, Anstee DJ, Cartron JP, Dahr W, Issitt PD, Jorgensen J, Kornstad L, Levene C, Lomasfrancis C, Lubenko A, Mallory D, Moulds JJ, Okubo Y, Overbeeke M, Reid ME, Rouger P, Seidl S, Sistonen P, Wendel S, Woodfiel G, Zelinski T (1995) *Vox Sang* 69:265
112. Stamatoyannopoulos G, Majerus PW, Perlmutter RM, Varmus H (2000) *The Molecular Basis of Blood Diseases*, 3rd ed. W. B. Saunders, Orlando, Florida
113. Morgan WTJ, Watkins WM (2000) *Glycoconjugate J* 17:501
114. Clausen H, Hakomori S-I (1989) *Vox Sang* 56:1
115. Hakomori S-I, Zhang Y (1997) *Chem Biol* 4:97
116. Keding SJ, Danishefsky SJ (2003) Synthetic Carbohydrate-Based Vaccines. In: Wong C-H (ed) *Carbohydrate-Based Drug Discovery*, vol 1. Wiley-VCH, Weinheim, p 381
117. Warren J, Geng X, Danishefsky S (2007) *Top Curr Chem* 267:109
118. Leloir LF (1971) *Science* 172:1299
119. Brockhausen I, Schachter H (1997) Glycosyltransferases Involved in *N*- and *O*-Glycan Biosynthesis. In: Gabius H-J, Gabius S (eds) *Glycosciences: Status and Perspectives*. Chapman & Hall, Weinheim, p 79
120. Kornfeld R, Kornfeld S (1985) *Annu Rev Biochem* 54:631
121. Ünlügil UM, Rini JM (2000) *Curr Opin Struct Biol* 10:510
122. Hu Y, Walker S (2002) *Chem Biol* 9:1287
123. Ramakrishnan B, Qasba PK (2001) *J Mol Biol* 310:205
124. Koeller KM, Wong C-H (2000) *Chem Rev* 100:4465
125. Trincon A, Giordano A (2006) *Curr Org Chem* 10:1163
126. Wong CH, Haynie SL, Whitesides GM (1982) *J Org Chem* 47:5416
127. Unverzagt C, Kunz H, Paulson JC (1990) *J Am Chem Soc* 112:9308
128. Henrissat B, Davies G (1997) *Curr Opin Struct Biol* 7:637
129. Wolfenden R, Lu X, Young G (1998) *J Am Chem Soc* 120:6814
130. Koshland DE (1953) *Biol Rev* 28:416
131. Zechel DL, Withers SG (2001) *Curr Opin Chem Biol* 5:643
132. Zechel DL, Withers SG (2000) *Acc Chem Res* 33:11
133. Vasella A, Davies GJ, Böhm M (2002) *Curr Opin Chem Biol* 6:619
134. Heightman TD, Vasella AT (1999) *Angew Chem* 111:794
135. Lillelund VH, Jensen HH, Liang X, Bols M (2002) *Chem Rev* 102:515
136. von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dyason JC, Jin B, Van Phan T, Smythe ML, White HF, Oliver SW, Colman PM, Varghese JN, Ryan DM, Woods JM, Bethell RC, Hotham VJ, Cameron JM, Penn CR (1993) *Nature* 363:418
137. McClellan K, Perry CM (2001) *Drugs* 61:261
138. Crout DHG, Vic G (1998) *Curr Opin Chem Biol* 2:98
139. Mackenzie LF, Wang Q, Warren RAJ, Withers SG (1998) *J Am Chem Soc* 120:5583
140. Williams SJ, Withers SG (2002) *Aust J Chem* 55:3
141. Hancock SM, Vaughan MD, Withers SG (2006) *Curr Opin Chem Biol* 10:509
142. Perugini G, Cobucci-Ponzano B, Rossi M, Moracci M (2005) *Adv Synth Catal* 347:941
143. Drickamer K, Taylor ME (1993) *Annu Rev Cell Biol* 9:237
144. Rini JM (1995) *Annu Rev Biophys Biomol Struct* 24:551
145. Weis WI, Drickamer K (1996) *Annu Rev Biochem* 65:441
146. Vijayan M, Chandra N (1999) *Curr Opin Struct Biol* 9:707
147. Drickamer K (1988) *J Biol Chem* 263:9557
148. Drickamer K (1995) *Nat Struct Biol* 2:437
149. Lee RT, Lee YC (2000) *Glycoconjugate J* 17:543
150. Mammen M, Choi S-K, Whitesides GM (1998) *Angew Chem Int Ed* 37:2754
151. Lundquist JJ, Toone EJ (2002) *Chem Rev* 102:555
152. Rini JM, Lobsanov YD (1999) *Curr Opin Struct Biol* 9:578
153. Crocker PR, Feizi T (1996) *Curr Opin Struct Biol* 6:679
154. Collins BE, Paulson JC (2004) *Curr Opin Chem Biol* 8:617
155. Schrag JD, Procopio DO, Cygler M, Thomas DY, Bergeron JJM (2003) *Trends Biochem Sci* 28:49
156. Moremen KW, Molinari M (2006) *Curr Opin Struct Biol* 16:592
157. Mesaeri N, Nakamura K, Zvaritch E, Dickie P, Dziak E, Krause KH, Opas M, MacLennan DH, Michalak M (1999) *J Cell Biol* 144:857

158. Van den Hamer CJA, Morell AG, Scheinberg IH, Hickman J, Ashwell G (1970) *J Biol Chem* 245:4397
159. Morell AG, Gregoriadis G, Scheinberg IH, Hickman J, Ashwell G (1971) *J Biol Chem* 246:1461
160. Ashwell G, Harford J (1982) *Annu Rev Biochem* 51:531
161. Meier M, Bider MD, Malashkevich VN, Spiess M, Burkhard P (2000) *J Mol Biol* 300:857
162. Ashwell G (1994) The physiological role of the hepatic lectin for asialoglycoproteins: An evaluation of current hypotheses. In: Beuth J, Pulverer G (eds) *Lectin Blocking: New Strategies for the Prevention and Therapy of Tumor Metastasis and Infectious Diseases (Zentralbl. Bakteri-ol., Suppl. 25)*. Gustav Fischer Verlag, Stuttgart, p 26
163. Hooper LV, Manzella SM, Baenziger JU (1997) The Biology of Sulfated Oligosaccharides. In: Gabius H-J, Gabius S (eds) *Glycosciences: Status and Perspectives*. Chapman & Hall, Weinheim, p 261
164. Kornfeld S (1990) *Biochem Soc Trans* 18: 367
165. Roberts DL, Weix DJ, Dahms NM, Kim J-JP (1998) *Cell* 93:639
166. Dahms NM, Hancock MK (2002) *Biochim Biophys Acta* 1572:317
167. Springer TA (1995) *Annu Rev Physiol* 57:827
168. Lasky LA (1995) *Annu Rev Biochem* 64:113
169. Kansas GS (1996) *Blood* 88:3259
170. McEver RP (2002) *Curr Opin Cell Biol* 14:581
171. Poppe L, Brown GS, Philo JS, Nikrad PV, Shah BH (1997) *J Am Chem Soc* 119:1727
172. Somers WS, Tang J, Shaw GD, Camphausen RT (2000) *Cell* 103:467
173. Newman W, Mirabelli CK (1998) *Expert Opin Invest Drugs* 7:19
174. Bendas G (1999) *Pharm Ztg* 144:3754
175. Wardlaw AJ (1999) *Drugs Future* 24:279
176. Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) *Chem Rev* 98:833
177. Kaila N, Thomas BE, IV (2002) *Med Res Rev* 22:566
178. Ernst B, Kolb HC, Schwardt O (2006) *Carbohydrate Mimetics in Drug Discovery*. In: Levy DE, Fügedi P (eds) *The Organic Chemistry of Sugars*. CRC Press, Boca Raton, FL, p 803