Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

Review Article

The sugar code: letters and vocabulary, writers, editors and readers and biosignificance of functional glycan–lectin pairing

Herbert Kaltner[1], José Abad-Rodríguez[2], Anthony P. Corfield[3], Jürgen Kopitz[4] and Hans-Joachim Gabius[1]

[1]Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University Munich, Veterinaerstr. 13, 80539 Munich, Germany; [2]Membrane Biology and Axonal Repair Laboratory, Hospital Nacional de Parapléjicos (SESCAM), Finca La Peraleda s/n, 45071 Toledo, Spain; [3]School of Clinical Sciences, Bristol Royal Infirmatory, Mucin Research Group, University of Bristol, Bristol BS2 8HW, U.K.; [4]Institute of Pathology, Department of Applied Tumor Biology, Faculty of Medicine, Ruprecht-Karls-University Heidelberg, Im Neuenheimer Feld 224, 69120 Heidelberg, Germany

Correspondence: Jürgen Kopitz (Juergen.kopitz@med.uni-heidelberg.de) or Hans-Joachim Gabius (gabius@tiph.vetmed.uni-muenchen.de or gabius@lectins.de)

Ubiquitous occurrence in Nature, abundant presence at strategically important places such as the cell surface and dynamic shifts in their profile by diverse molecular switches qualifies the glycans to serve as versatile biochemical signals. However, their exceptional structural complexity often prevents one noting how simple the rules of objective-driven assembly of glycan-encoded messages are. This review is intended to provide a tutorial for a broad readership. The principles of why carbohydrates meet all demands to be the coding section of an information transfer system, and this at unsurpassed high density, are explained. Despite appearing to be a random assortment of sugars and their substitutions, seemingly subtle structural variations in glycan chains by a sophisticated enzymatic machinery have emerged to account for their specific biological meaning. Acting as 'readers' of glycan-encoded information, carbohydrate-specific receptors (lectins) are a means to turn the glycans' potential to serve as signals into a multitude of (patho)physiologically relevant responses. Once the far-reaching significance of this type of functional pairing has become clear, the various modes of spatial presentation of glycans and of carbohydrate recognition domains in lectins can be explored and rationalized. These discoveries are continuously revealing the intricacies of mutually adaptable routes to achieve essential selectivity and specificity. Equipped with these insights, readers will gain a fundamental understanding why carbohydrates form the third alphabet of life, joining the ranks of nucleotides and amino acids, and will also become aware of the importance of cellular communication via glycan–lectin recognition.

Introduction

Clearly, molecular languages of life govern cellular communication. The long and winding road that the concept of information coding by glycans has taken to gain acceptance is first surveyed by the selected quotations from the literature. They graphically emphasize that two classes of biomolecules have overshadowed glycans for decades. 'To an observer trying to obtain a bird's eye view of the present state of biochemistry, …. life may until very recently have seemed to depend on only two classes of compounds: nucleic acids and protein' [1].

As is reflected in the manner of covering the subject of life sciences in textbooks, the enormous technological advances in the study of these two classes of biomolecules turning sequence and expression analyses into routine practice were a reason to maintain this focus. In this sense logically, 'with the rapid advent of powerful insights and tools, the study of nucleic acids and proteins came to dominate the molecular biology revolution of subsequent decades' [2]. When analyzing the flow of biological information with the aim of reaching an understanding of the molecular basis of (patho)

Received: 11 July 2019
Revised: 31 August 2019
Accepted: 4 September 2019

Version of Record published:
24 September 2019

Biochemical Journal (2019) 476 2623–2655
https://doi.org/10.1042/BCJ20170853

physiological processes, however, the potential of carbohydrates to serve as an alphabet for 'writing messages' should not be definitely overlooked. Especially on cell surfaces, where space is limited, signals in the form of oligomers must meet the requirement of being compact and precise. In chemical terms, individual 'letters' (monomers) must be endowed with the capacity to build a rich vocabulary (oligomers) that can be 'read' to then exert roles in cellular communication.

Ubiquitous in Nature, even constituting the most abundant compounds on the Earth, namely the polymers, cellulose and chitin [3], and also prominently positioned on cell surfaces [4], carbohydrates qualify as more than being the biochemical fuel and molecular concrete of cell walls. Terminologically, the common presence of glycans as part of glycoproteins and glycosphingolipids was an incentive for coining the term 'glycobiology' [5] (or more general: glycosciences). It implies fundamental functionality. Conceptually, the term gave research work on glycans a clear direction. As a consequence, carbohydrates, like amino acids and nucleotides, gained the status of an alphabet of life [6]: 'letters' of this alphabet are shown in Figure 1. Chemical modifications, such as the derivatization of glucose (Glc) or galactose (Gal) to their 2-amino sugars and then to *N*-acetylated products (GlcNAc, GalNAc), can be likened to Umlaut formation in the German language. This process is especially frequent in the case of *N*-acetylneuraminic acid (Figure 1, bottom row — center). Its modification yields more than 80 different sialic acids, in mammals prominently *O*-acetylation and hydroxylation to *N*-glycolylneuraminic acid [7,8]. Thus, the 'letters' of an alphabet are available. By having a close look at the meaning of the term 'carbohydrates' and the structures shown in Figure 1, the fundamental potential of sugars to store (in 'words') and to transfer biological information (by molecular recognition) will become evident.

Carbohydrates: the third alphabet of life

The name 'carbohydrates' originates from the stoichiometric proportion of carbon and water ($C_n(H_2O)_m$, with $n \geq m$). It reflects that these compounds share the presence of hydroxyl groups. The chemical equivalence of most of them opens the door to an unsurpassed level of structural diversity that glycans attain beyond the sequence. Ironically, precisely this special chemical characteristic to make an exceptional number of isomers possible in a minimum of chain length (reasons for this will be given below) was responsible for the slow rate of progress in this field, here highlighted in a Q & A manner as follows: 'in this remarkable age of genomics, proteomics and functional proteomics, I am often asked by my colleagues why glycobiology has apparently so far lagged behind the other fields. The simple answer is that glycoconjugates are much more complex, variegated, and difficult to study than proteins or nucleic acids' [9]. Figure 2 explains the chemical basis to enable structural diversity.

The building blocks of nucleic acids and of proteins are linked in a uniform manner, the 5′,3′-phosphodiester and the peptide bonds (Figure 2). Obviously, the chemical nature of the glycosidic linkage, too, is the same throughout an oligo- or polymer. In complete contrast, many sites, that is each hydroxyl group of the sugars shown in Figure 1, can, in principle, serve as acceptor during chain elongation. This new level of versatility is symbolized by the arrows in Figure 2 that point to these functional groups. If this happens, the resulting products will be isomers. Their sequence can then no longer be entirely defined by the order of the building blocks, as is the case for nucleic acids or peptides, and there is another source of variability. Forming a glycosidic bond necessarily involves either the α- or the β-anomeric position of the donor, i.e. the activated sugar, which is highlighted by the pair of arrows pointing away from the pyranose (Figure 2).

Bringing these two factors for structural variability together, clearly more than a single compound will satisfy the definition of a diglucoside. Using the diglucoside (Glc–Glc) as an example, the permutations, indicated by arrows in Figure 2, are possible. Both the anomeric position and the linkage points come into play, not just a single structure (as is the case for a dinucleotide or a dipeptide) but 11 diglucosides (eight diglucosides with an acceptor site for chain elongation) become possible. As a consequence, beyond the sequence, the type of anomery (α or β) and the linkage position from the anomeric center at the C1 atom of the donor to the acceptor (for hexopyranoses: 1,1; 1,2; 1,3; 1,4; 1,6 as given in Figure 2) must therefore be known to fully define each disaccharide. This explanation clearly substantiates why the structural analysis of carbohydrates is much more difficult than that of nucleic acids or proteins. Factors for further increasing the pool size of oligo- and polysaccharides can include the ring size (common hexasaccharides, such as Gal, are present as pyranose, as given in Figure 1, but Gal is also a part of polysaccharides and glycoconjugates of bacteria, fungi and protozoa as furanose (Gal*f*) [10–12]) and the implementation of site-specific substitutions, the latter being the biochemical equivalent of cytosine methylation/hydroxymethylation or posttranslational protein modifications [13]. Phosphorylation (for example of mannose at its 6-position) or *O*-sulfation (for example of Gal or GalNAc at

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

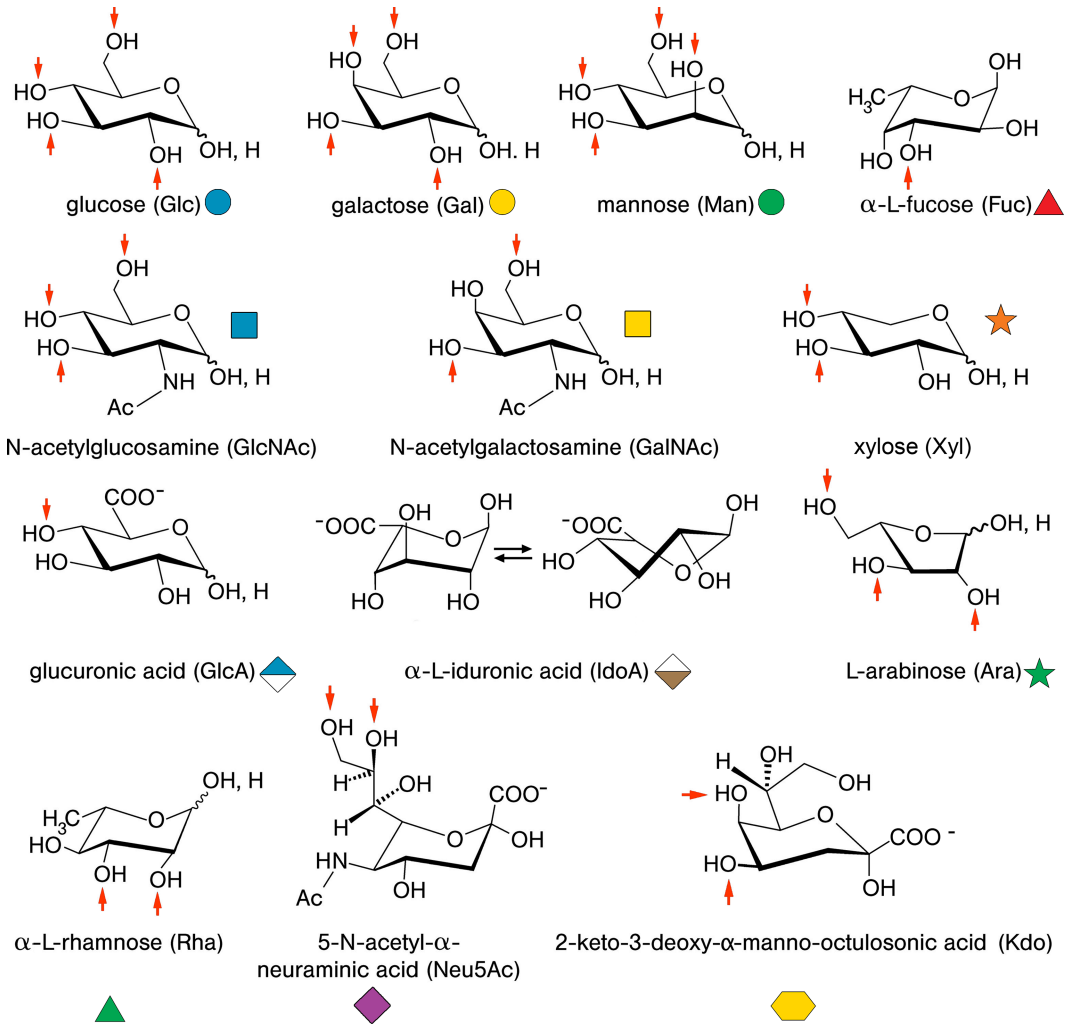


Figure 1. Letters of the sugar language.

Illustration of the main letters of the third alphabet of life. In each case, the structure, name and symbol as well as known acceptor positions (by arrows) in glycoconjugates are presented. Four sugars have L-configuration: fucose (6-deoxy-L-galactose), rhamnose (6-deoxy-L-mannose) and arabinose are introduced during chain elongation or at branch positions, whereas L-iduronic acid (IdoA) results from the enzymatic epimerization of GlcA (at C5) in glycosaminoglycans that is sandwiched between the *N*-deacetylase/*N*-sulfotransferase and the *O*-sulfotransferase reactions. Its $^{1}C_{4}$ conformer (left) is in dynamic equilibrium with the $^{2}S_{0}$ form (right) that has the appropriate stereochemical arrangement for glycosaminoglycan (heparin/heparan sulfate)-receptor binding (please see also Figure 3d). Xylose is part of the GlcA-$(Gal)_2$-Xyl linker between the proteoglycan core protein and the disaccharide repeat section of *O*-linked glycosaminoglycans (chondroitin, dermatan and heparan sulfates). Neu5Ac, one of more than 80 sialic acids, is a common constituent of glycoprotein and ganglioside glycans. Its presence and the α2-linkage to 3-, 6- or 8-positions of acceptors give the respective oligosaccharide a specific meaning. Kdo is a building block of lipopolysaccharides in the cell wall of Gram-negative bacteria and is also found in cell wall polysaccharides of green algae and higher plants. Here, polysaccharides, such as arabinoxylans or pectins (such as rhamnogalacturonans), are abundant (from [248], extended; with permission).

3- or 4-positions), as illustrated in Figure 3, gives the respective letter a new particular meaning (for respective information, please see legend to this figure and also below).

Succinctly summarized in direct comparison to nucleic acids and proteins, 'carbohydrates are ideal for generating compact units with explicit informational properties since' (as graphically illustrated in Figure 2) 'the permutations on linkages are larger than can be achieved by amino acids and, uniquely in biological polymers, branching is possible' [14]. A branched pentasaccharide is shown in Figure 4. Elegant acceptor–donor tailoring

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

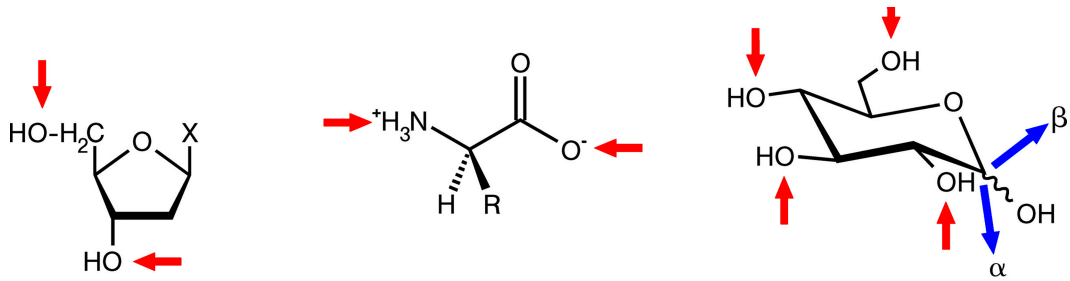


Figure 2. How letters of the three alphabets of life form words.

Illustration of the linkage points between building blocks in biopolymers. The comparison highlights the fundamental difference between the uniform backbone structures of nucleic acids and proteins and the variability within chain elongation of glycans of cellular glycoconjugates. The red arrows symbolize the 5′,3′-phosphodiester between nucleotides and the peptide bonding involving the two functional groups of the C_α-atom that all amino acids have, i.e. amino and carboxyl groups. In contrast with these pairs connected at fixed positions, each of the chemically equivalent hydroxyl groups can, in principle, be used as an acceptor by glycosyltransferases. They transfer the monosaccharide from the activated donor in α- or β-anomeric position to build a chain (for example of two disaccharides and a pentasaccharide, please see Figure 3b–d; for glycolipid glycans, please see Table 2; from [248], with permission).

has made chemical glycan synthesis possible [15], documenting the feasibility to prepare isomers as distinct compounds in the laboratory. The price to be paid at the cellular level toward this end will be to devote a share of the genome to coding for glycogenes, whose products will then make biochemical assembly possible.

In Nature, a complex enzymatic apparatus has been developed to realize a portion of the theoretical potential for diversity using carbohydrates as substrates [16–23]. It works in the cell at different sites including the nucleus and also on the cell surface. The enormously high-degree spatial organization of glycosyltransferases at a center of glycosylation along the route to secretion, i.e. the Golgi apparatus, has intrigued researchers for decades [24,25]. Sophisticated analytical technology applied to the products of the assembly lines has accomplished mapping of structural diversity based on the mentioned permutations [26–28]. Indeed, reaching the full range of linkage positions, for example in the case of adding the 'letter' fucose (see Figure 1) with its α1,2, α1,3, α1,4 and α1,6 connections [20], is documented; and the enzymatic machinery is even able to tailor the local density of glycans by different means.

A key prerequisite toward this end is the diversification within a glycosyltransferase family. To give examples already at this point, the degree of branching of N- and (mucin-type) *O*-glycans is varied due to the existence of panels of *N*-acetylglucosaminyltransferases (GnTs) so that the number of antennae and thus of terminal epitopes in close vicinity is adaptable [16]. Fittingly, an engineered loss of such an enzyme responsible for branching appears to be functionally compensated *in vivo* by shifts in the glycome including neosynthesis of glycan chains (by *O*-mannosylation), homeostasis of elaborated branch ends and extension of chains by including *N*-acetyllactosamine (LacNAc) repeats [21,29–33]. The same principle is operative at the level of initiation of protein glycosylation. The large size of the toolbox for starting mucin-type *O*-glycosylation by UDP-GalNAc: polypeptide *N*-acetylgalactosaminyltransferases (GalNAcT) and their broad spectrum of (glyco)peptide specificity provide an instructive example for modulating the density of separate glycans, the human genome encoding 20 members of this family [34–36]. As noted above, subsequent chain elongation can then include branching so that the full spectrum from sparse to highly clustered glycosylation (at different sites and with different degrees of branching) is covered.

This way, and with elaborate branch-end tailoring, the enzymatic machinery (for example of the mentioned *O*-glycan [37–39]) decorates cell surfaces with glycan-based signals of a highly sophisticated structure and a varying degree of local density. These sets of enzymatic reactions, which will be explained below, add up to let the molecular diversity and thus the coding capacity of glycans attain the top position among biopolymers. Since the presence, localization and activity of each component of the glycan-producing factory are intimately controllable, a cell can express, for example, its status of differentiation by distinct glycans such as stage-specific embryonic antigens (SSEA) [40–42]. They are often defined as being the targets of monoclonal antibodies and have been given a place in the CD nomenclature such as CD15 (SSEA-1, Lewis^x) [43].

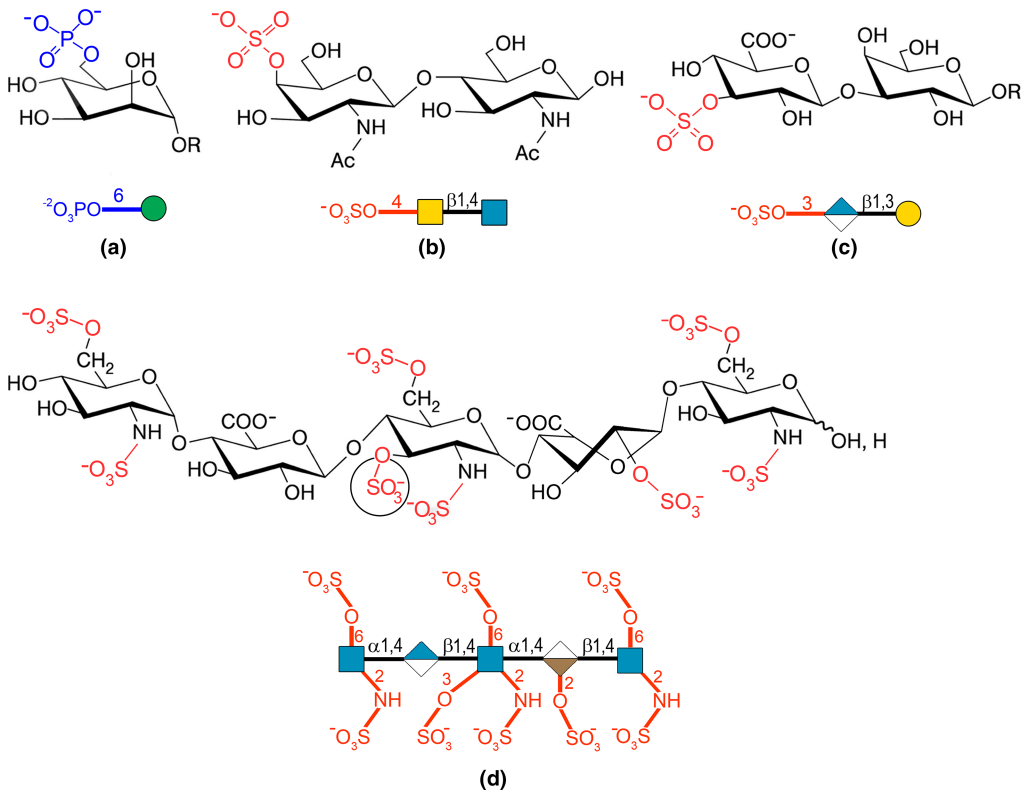


Figure 3. Sugar-based signals.

Illustration of (sugar code) 'words' (signals) with biomedical significance. (**a**) Man-6-phosphate: enzymatic labeling of distinct Man moieties in high-mannose- or hybrid-type *N*-glycans of distinct glycoproteins by 6-phosphorylation in a 2-step reaction (using first UDP-GlcNAc as the substrate for incorporation of a GlcNAc-1-phosphate and then removal of the GlcNAc unit by a phosphodiester glycosidase) in the *cis*-Golgi that facilitates binding of the resulting Man-6-phosphate moiety to the two P-type lectins in the *trans*-Golgi network and transport of, for example, lysosomal hydrolases; (**b**) 4-*O*-sulfated LacdiNAc: this type of *N*-glycan terminus labels distinct glycoproteins, such as pituitary glycoprotein hormones for lectin-mediated uptake by hepatic endothelial cells; (**c**) HNK (human natural killer cell)-1 epitope (CD57): it is involved in cell adhesion/migration in the nervous system; (**d**) heparin pentasaccharide: the *N*,*O*-sulfation pattern with especially the encircled 3-*O*-sulfate group in the trisubstituted central GlcN unit (sulfation code) accounts for potent anti-coagulant activity. Considering the conformer conversion of IdoA to its $^{2}S_{0}$ form shown in Figure 1, 2-sulfation of IdoA and sulfation in its vicinity favors adaption of the skew-boat structure and thus intimate binding to antithrombin III [249,250] (from [248], modified and extended; with permission).

In general terms, glycan production is not only fundamentally different from polymerase- and ribosome-dependent nucleic acid and protein synthesis by its in-built variability. It is also template independent, the proteins encoded by glycogenes acting in concerted but versatile action. The occurrence of the noted variability is, therefore, not governed by a strict master plan but is implemented via levels of substrate availability (donors and acceptors), transporters and enzymes in the local context, which allows highly dynamic changes to occur swiftly. Moreover, the evidently spatiotemporally highly regulated and thus non-random remodeling by enzymatically shortening glycans is known (please see below for details). These factors add up to recognize why the cellular glycome can be likened to a molecular fingerprint. Evidently, the virtues of carbohydrates for structural variability are used in a far-reaching manner, a strong argument for a broad functional significance of glycans.

Moving on from these two structural dimensions (of linear and branched glycans), the next question is whether analyzing the third dimension, i.e. glycan shape and its dynamics, may also have something in store for a role in information handling. Indeed, restrictions to flexibility for rotations of the bulky sugars about the dihedral angles of each glycosidic bond are often seen [44,45]. As a consequence, each combination of this

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

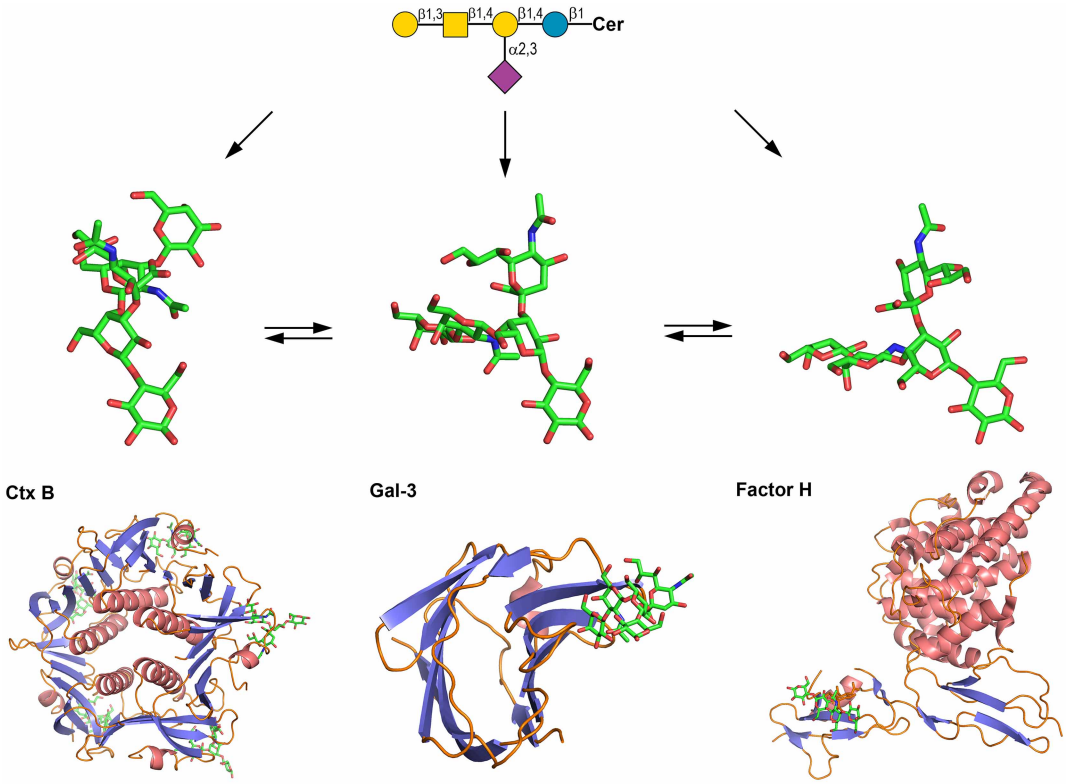


Figure 4. Conformer selection of a glycan ligand by receptor binding.

Illustration of the structure of ganglioside GM1 with its pentasaccharide (top), three low-energy conformers (center) and the crystallographical structure of a complex of each conformer with a receptor (bottom). The cases of the pentameric lectin part of cholera toxin (3CHB), human galectin-3 (3AYC) and human factor H (4ZH1) are presented.

parameter set (the Φ, Ψ-angles that define the shape of a disaccharide; the Φ, Ψ, ω-angles when a third bond, such as that between C5 and C6, is involved in the connection between two sugar units, as is the case in a 1,6-linkage) will define a certain shape, and each one will have its own (favorable or unfavorable) energy value. The conformationally accessible space will be hereby structured into regions differing in energy levels like a topographical map, low-energy levels representing valleys in a hilly landscape [46]. A glycan can access these regions so that it 'moves in solution through a bunch of shapes' (conformers or molecular keys) [47], which is the distinctive feature of the third dimension of the sugar code.

This preferential occurrence of distinct conformers has remarkable implications for molecular interactions: it reduces the entropic penalty within the thermodynamics of a binding process when involving a certain conformer that docks into a site of complementarity — in comparison with a highly flexible ligand. What necessitates folded surroundings or disulfide bridges in a protein to force a peptide fragment to adopt a certain conformation comes by Nature in oligosaccharides. How different, energetically favored conformers can look like is illustrated for a branched pentasaccharide of a glycosphingolipid, i.e. ganglioside GM1, in Figure 4.

The concept of the sugar code stipulates that such dynamic shape diversity will have consequences: the distinct spatial positioning of the free hydroxyls (or also an *N*-acetyl group) creates a particular surface in each conformer like a stretched-out hand. Complementarity with a receptor engenders hydrogen and coordination bonding as well as by C-H/π-interactions between slightly positively polarized C-H groups (of the B face of Gal) and π-electrons of tryptophan, a reason why this rarely occurring amino acid is often present in the contact sites for sugars [48–51]. Of note, a change in the position of a single hydroxyl group, for example to present the hydroxyl group in position 4 of hexopyranoses axially (please compare Glc to its epimer Gal in Figure 1), is sufficient to give the letter a new meaning. Each hydroxyl group in a carbohydrate can, therefore, have a mission in connecting 'letters' to 'words' or, alternatively, in enabling 'words' to be distinguishable.

It, thus, was prophetic to envision a sugar code, in one occasion inspired by the local vicinity of labs: 'drawing an analogy to the genetic code, [Victor] Ginsburg, a biochemist whose laboratory was down the hall [close to the Marshall Nirenberg laboratory where the genetic code was being cracked at that time], was convinced that after the genetic code was deciphered, the next important code to solve was the one for cellular recognition; and to him it seemed likely that this was based on specific protein-glycoconjugate interactions … this seemed like an idea worth pursuing at a future time' [52]. About the same topic, Albert Neuberger, a pioneer of glycoprotein research, went one step further by succinctly stating that 'it is almost certain that glycoproteins are prominently involved with social behaviour of cells' so that it is 'reasonable to believe that this complicated mechanism' [of glycan biosynthesis] 'serves an important biological function' [53], and conveying a molecular message by glycans has turned out to be the case. That providing convincing support for the idea that glycans are on par with proteins and nucleic acids for the purpose of information coding, the basis of the concept of the sugar code [54], today is a highly active and dynamic research area has recently been underscored by the comment that 'only in recent years have we begun to appreciate how deeply glycan functions pervade all aspects of organismic biology, molecular biology, and biochemistry' [55].

This said, it is clear that there is a need for an introduction to the functionality of glycans/glycoconjugates based on the selective binding of glycans ('reading' glycan-encoded messages) by carbohydrate–protein recognition. Having just equipped our readership with the knowledge on why carbohydrates are predestined as 'letters' for 'writing' molecular messages so that their oligomers become 'words' with high-coding capacity, the next task is to explain why their presence in cellular glycoconjugates can fulfill the mentioned promises. To do so is to take the term 'reading a glycan-encoded message' literally and to introduce a tool that is able to choose, pick, read or select (Latin: *legere*). This etymological root explains the origin of the term 'lectin' for (glyco)proteins found in plants (and invertebrates) that have an antibody-like selectivity enabling them to distinguish and to agglutinate erythrocytes of different donors according to their histo-blood group ABH status, thus also called phytohaemagglutinins [56]. The inherent epitope specificity and the nature of the bound determinant as sugar was the reason why this term became popular for carbohydrate-binding proteins, excluding immunoglobulins, enzymes and sensor/transport proteins for free mono- to oligosaccharides, [50,57–61]. Their ubiquitous presence on the Earth matches the postulate of their importance. The members of this class of proteins are thus able to engage in pairing with cognate glycans, a talent useful for glycophenotyping of cells and tissues [62–65], and this interaction should then trigger post-binding effects, making the pairing functional. To do so, lectins are expected to fulfill a series of structural prerequisites.

Lectins: readers of sugar-encoded information and effectors

In principle, the high level of glycan diversity should find its equivalent on the side of the lectins in order to achieve a broad functional spectrum for glycan–protein recognition. Fittingly, more than a dozen folds have acquired the capacity to bind carbohydrates (for listing, please see Table 1; for illustration of the folds, please see Galleries of Lectins [51,66]; for a review on structures, please see [67]). Obviously, developing the structural basis for lectin activity has not been a rare event. Supporting this reasoning, a certain basic fold can even serve as a platform to generate several types of carbohydrate recognition domains (CRDs) with their own characteristic properties. This principle is illustrated in Figure 5 for the β-sandwich as structural unit that acquires very specialized binding properties by independent routes (please see legend to this figure for further information and also Table 1). The positions of the contact site of sugar within the fold, its sequence signature and three-dimensional structure as well as the type of involvement of a Ca^{2+} differ between the individual group members, reflecting that the course of evolution has taken separate roads toward these distinct phenotypes, which share the same origin. Clearly, structural CRD diversity lives up to the expectation to match glycan diversity, and this finding immediately leads to the question whether this also holds true for the number of glycans that have become ligands for lectins: it should be large.

As the current status of research reveals, the inventory of tissue lectins can indeed bind to a wide panel of glycans. As the first class of targets with clinical relevance, foreign glycan epitopes are readily detected by tissue lectins as part of the innate host defense. Instructive examples illustrate the efficiency of this kind of surveillance mechanism independent of antibodies: tracing the mentioned Gal*f* (by human intelectin-1 [68]), the bacterial cell wall peptidoglycan with the β1,4-linked disaccharide of GlcNAc and *N*-acetylmuramic acid, the substrate of lysozyme, (by bactericidal C-type lectin RegIIIγ (murine)/*h*epato*i*ntestinal *p*ancreatic or

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

Table 1 Classification of animal/human lectins by folds

Type of fold	Example for lectin
C-type	Asialoglycoprotein receptors (hepatocytes, macrophages), collectins, selectins, tandem-repeat-type receptors
I-type	Siglecs and structurally related paired receptors
P-type	Ca^{2+}-(in)dependent mannose-6-phosphate receptors (MR) and proteins with MR homology domain such as erlectin
β-sandwich	Molecular chaperones (calnexin, calreticulin) ER-to-Golgi glycoprotein transporters such as ERGIC-53 or VIP36 Module in ubiquitin ligase complex and *N*-glycanase Galectins Pentraxins Laminin G-like CRD, e.g. for matriglycan binding
β-trefoil	Cysteine-rich domain of tandem-repeat-type C-type lectins such as the macrophage mannose receptor (MMR, CD206) Lectin domain of GalNAc-Ts with activity at one (GalNAc-T2/4/10) or two (GalNAc-T1) of the three sites
β-propeller (5-bladed)	tachylectin-2
β-propeller (6-bladed)	tachylectin-1
β-propeller (7-bladed)	mushroom (*Psathyrella velutina*) agglutinin
β-prism I	zymogen granule protein-16 (ZG16p) and its paralogue ZG16b
β-prism II	pufferfish (*Fugu rubripes*) lectin
jelly-roll barrel	eel (*Anguilla anguilla*) agglutinin, *Xenopus* X-epilectin
TIM barrel	YKL-40 (chitinase-like lectin)
fibrinogen-like	ficolectins, intelectins
α+β	lectin domain of mouse latrophilin-1, a G-protein-coupled receptor
$(\alpha\alpha)_7$ barrel	ER-associated degradation-enhancing α-mannosidase-like proteins (EDEM)

*p*ancreatitis-*a*ssociated *p*rotein (HIP or PAP; human) [69]) or the mycobacterial cord factor trehalose-6,6′-dimycolate (by *m*acrophage *in*ducible *C*-type *le*ctin Mincle (also called CLEC4E or Clecsf9)) [70–72] and distinguishing glycan signatures typical for infectious agents such as a high local density of Man residues characteristic for the surfaces of yeast cells from those of normal cells by collectins, ficolins or tandem-repeat-type C-type lectins all reveal an involvement of lectins in protecting against infections [73–75]. In order to let the sugar code regulate (patho)physiological processes, glycans of cellular production presented by proteins or sphingolipids qualify as ligands, of course (please see below).

Equally important, when viewed from the perspective of the ligand, it is possible that different folds converge to target the same glycan determinant. This fundamental rule in glycan–lectin recognition is illustrated for the low-energy shapes of the pentasaccharide shown in Figure 4. Having realized the existence of the third dimension of the sugar code manifested by the presence of distinct glycan conformers, Figure 4 presents a proof-of-principle case of selection of each conformer by its specific lectin. The structural analysis of the respective glycan–lectin complexes has revealed that the full set of this glycan's energetically privileged conformers can serve as a ligand, and each lectin selects its cognate conformer (shape). Looking beyond this special case, not only different conformers but also the same shape can be accommodated by more than one fold. This is, for example, the case for the low-energy conformer of epitopes of branch-end epitopes of *N*- and *O*-glycans of glycoproteins such as LacNAc (please see below) [76,77].

Naturally, lectin generation did not stop at the stage of fold programming for sugar specificity. Using the gene for any ancestral CRD as a starting point, duplications were then instrumental in setting diversification in motion at two levels: (i) by introducing sequence deviations during evolution to let offspring CRDs gain their

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

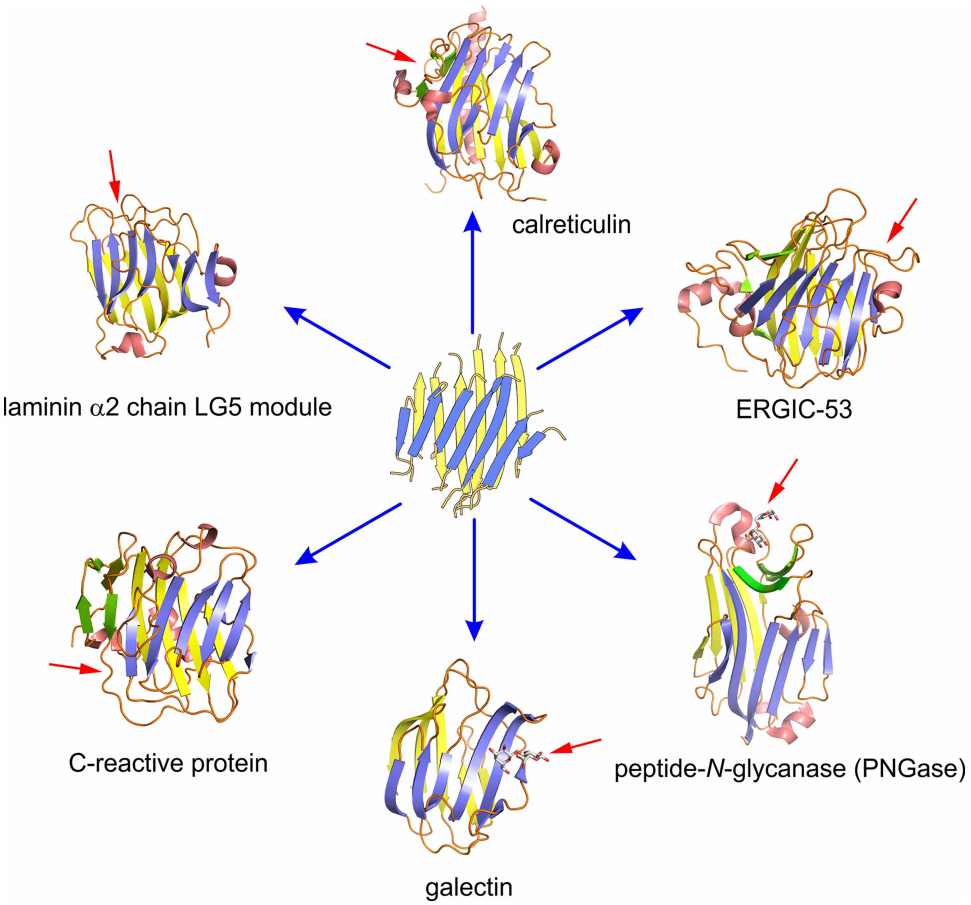


Figure 5. Multi-purpose use of a basic fold (i.e. β-sandwich) as lectin module.

Illustration of the β-sandwich fold (center) and its occurrence in molecular chaperones (murine calreticulin; 3O0X), ER-to-Golgi transporters (rat ERGIC-53; 1R1Z), *N*-glycanase (murine; 2I74), galectins (human galectin-1; 1GZW), pentraxins (human C-reactive protein; 3PVO) and the fifth laminin G-like module of mouse laminin α2 chain (1QU0). The arrow points to the site of contact with ligand.

own binding properties and (ii) by the mode of modular design, manipulating the quaternary structure or generating puzzle-like combinations with other (types of) domains. Focusing on the β-sandwich-type *ga*(lactose-binding)*lectins* as an instructive example, the route of evolution established a pool size of up to 15 proteins (in mammals), all with the same type of CRD (Figure 6). As postulated, the monitoring of the specificities of different galectins from the same organism, for example, by frontal affinity chromatography [78,79] or by glycan microarrays [80], confirmed the expectation for overlapping profiles with individual characteristics. Between species, not only do the fine-specificity features of galectins differ, but also species-specific characteristics in the gene pattern occur, as seen in Figure 6 in the mouse/rat comparison with a gain of a distinct lectin (galectin-5 or -6) in each case.

That such variations are rather common is further substantiated by the gene number variation for a C-type lectin of macrophages in human (one) and mouse (two) [81,82], even more pronounced for the cases of the *d*endritic *c*ell *i*nhibitory *r*eceptor (DCIR) with a single gene in human but four homologs in mouse and rat [83] and of another dendritic cell C-type lectin (i.e. DC-SIGN, also called CLEC4L or CD209) that is coded by one gene in human but eight genetic homologs are present in mice [84]. Single nucleotide polymorphisms (SNPs) take variability to the level of individuals of a species. The ensuing single-site substitutions, such as F19Y in human galectin-8 or K313I/N288D in a C-type lectin of Langerhans cells called langerin (CLEC5K, CD207), alter the structure and affect binding, for langerin increasing affinity to GlcNAc-terminated glycans and abolishing association with terminal 6-*O*-sulfated Gal, for example, presented by the glycosaminoglycan keratan sulfate [85,86].

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

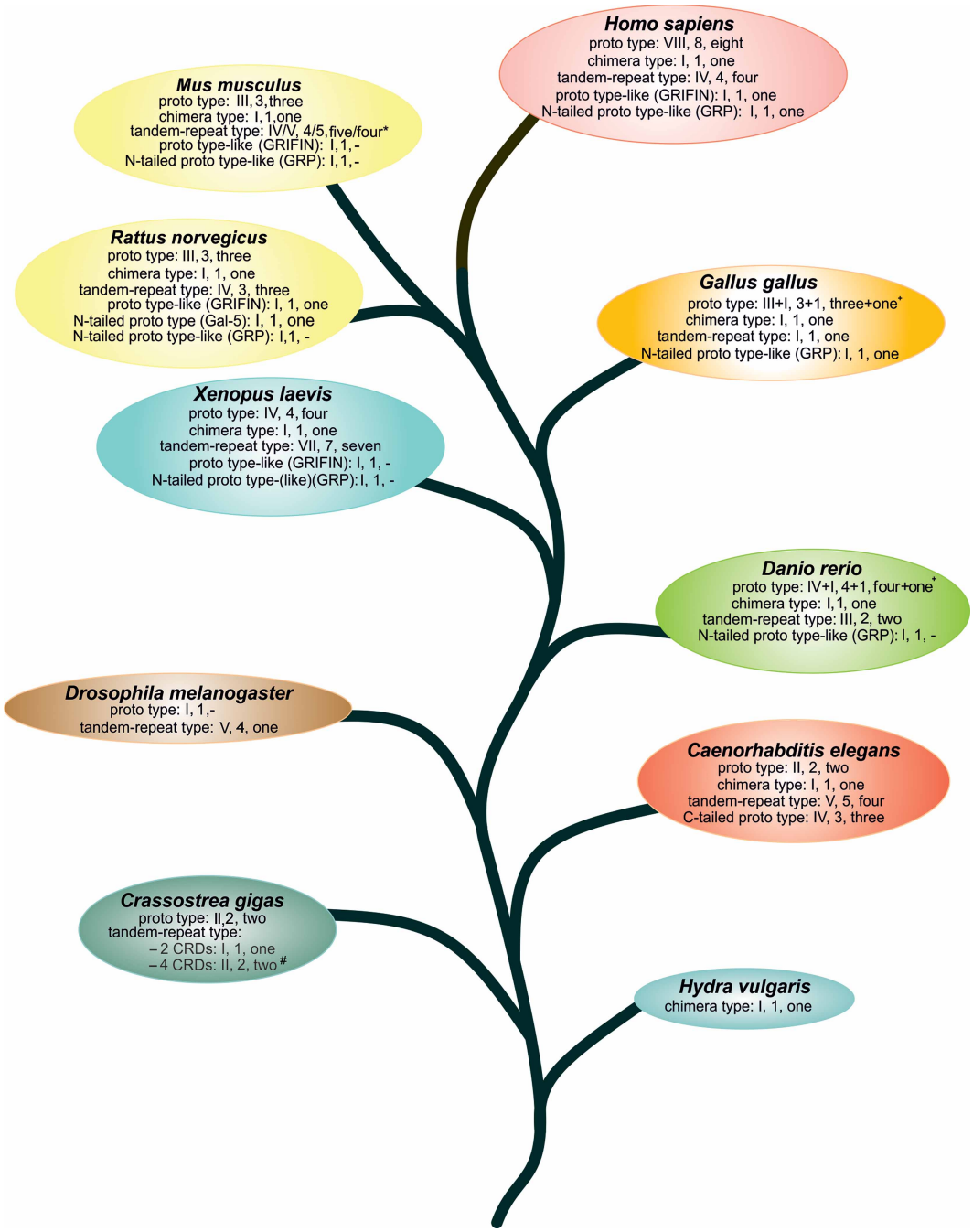


Figure 6. The galectin family in phylogenesis.

Illustration of the composition of the galectin family in model organisms. Evidence for galectin presence on the level of the gene (Roman number), the mRNA (Arabic number) and the protein (numerical information) is listed in each group. The given classification is based on the type of structural context of presentation of the canonical CRD (for its graphical depiction, please see Figure 7). *G*alectin-*r*elated *p*rotein (GRP) in vertebrates and galectins-7, -8, -10 and -11 in *C. elegans* harbor N- or C-terminal extensions, respectively. The proto-type galectin-5 (composed of a short N-terminal section attached to a sequence nearly identical with that of the C-terminal CRD of galectin-9) is found exclusively in rat, tandem-repeat-type galectin-6 exclusively in mouse, here subject to the presence/absence of polymorphism between strains (*). *G*alectin-*r*elated *i*nter-*f*iber prote*in* (GRIFIN) is special due to its species-dependent variability in lectin activity (+; loss in mammals, maintained in chicken and zebrafish). Two tandem-repeat-type galectins in the Pacific oysters (#) (and also Eastern/pearl oysters) are special as they have four different CRDs arranged as protein-like pearls on a string.

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

Having already mentioned the characteristics of the protein architecture beyond the basic fold and the sequence of the CRD above, the information presented in Figure 6 actually illustrates group building based on three types of spatial arrangement of the canonical CRD, classified by names. Structurally, they stand for (i) a non-covalently associated homodimer with identical CRDs (proto type), (ii) a linker-connected heterodimer joining two different CRDs (tandem-repeat type) and (iii) a modular combination of the CRD with an N-terminal region of non-triple helical collagen-like repeats (for self-aggregation) and two acceptor sites for serine phosphorylation called chimera type [87–90]. These three types of design of vertebrate galectins are depicted in Figure 7. Based on bringing CRDs together via different means, a common theme in lectin evolution with the highest level of diversification in the modular design achieved by C-type lectins [91], these proteins can serve assignments well to sense the presence of appropriate binding partners and then to bridge them, a hallmark of lectin activity — and there are many more functional read-outs than (hem)agglutination. An overview of the resulting activities of tissue lectins is given in Figure 8, depicting the wide range of influence of glycan–lectin recognition on (patho)physiological processes. Remarkably, modularity is a means of bringing the domains together that can functionally co-operate so that lectins acquire bi- and multifunctionality [59,92]. Targeting of enzyme action to a substrate by the CRD of the modular protein listed in Figure 8 is operative, for example, in initiating mucin-type *O*-glycan synthesis and quality control of glycoprotein folding (please see below).

Proceeding from this conceptual foundation, i.e. carbohydrates as platform to generate oligomers with high-density coding capacity in three dimensions and lectins as their receptors, the survey of routes of glycoconjugate synthesis that follows will immediately guide us to insights on how molecular switches for affinity, selectivity and specificity are installed. In essence, seemingly subtle structural differences will be shown to signify much more than, for example, the simple addition of a branch or the gain/loss of a sugar in a glycan. Considering the evidence for dynamic remodeling of glycans, the example presented with conspicuous significance for growth regulation adds to underscore the paradigmatic strength of the concept of the sugar code.

Glycoproteins: platform for presenting glycan-encoded information

Starting from the original observations that, in the author's words, it is not impossible that mucin 'einen gepaarten Stoff darstelle' [appears to be a paired substance], with protein and sugar components [93], and, decades later, that glycans are indeed covalently linked to protein, first shown for ovalbumin and a GlcNAc β1, N(Asn) linkage in ovalbumin [94], it became clear that a sugar part is an integral constituent of many proteins. On the surface of the plasma membrane, in the extracellular matrix and in serum, glycoproteins are abundantly present. By compiling information on the occurrence and properties of the glycan-protein linkage, it became

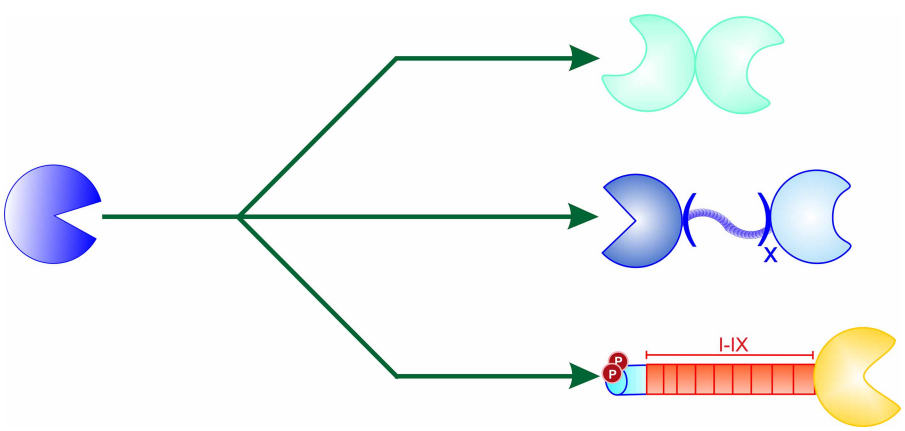


Figure 7. Structural design of vertebrate galectins.
Illustration of the three types of presentation of the canonical CRD in vertebrate galectins. It can form non-covalently associated homodimers (proto type), linker-connected heterodimers (tandem-repeat type) and a trimodular chimera (galectin-3) composed of an N-terminal section with two sites for Ser phosphorylation, the middle part with non-triple helical collagen-like repeats (nine in the human protein) and the C-terminal CRD.

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

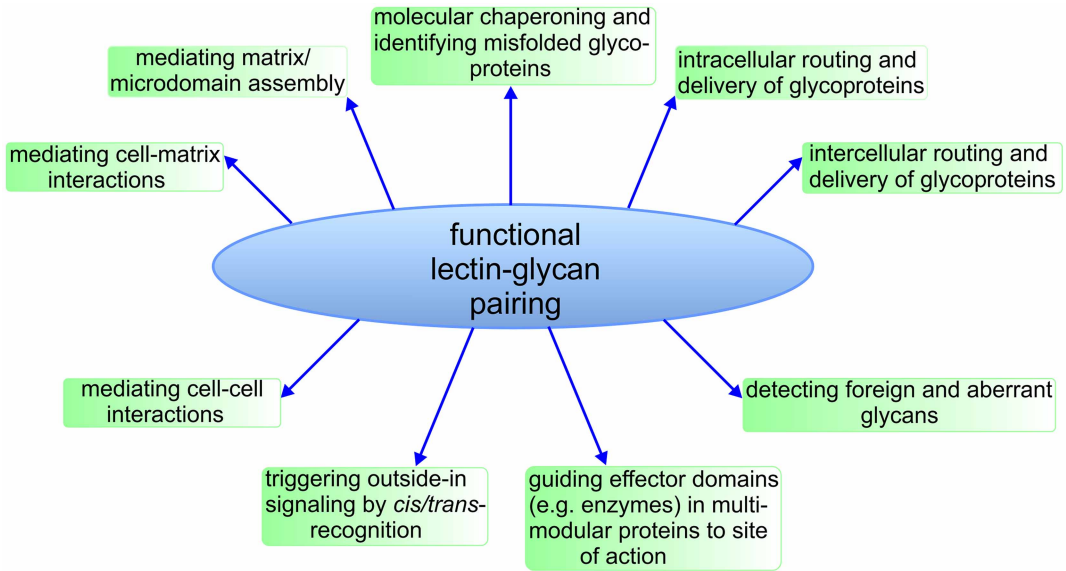


Figure 8. What lectin–glycan pairing can do.

Illustration of the range of functions exerted by the functional pairing of lectins with glycans of cellular glycoconjugates and foreign glycan signatures.

clear that 'the glycopeptide bond has attained the broadest possible phylogenetic distribution', with 13 monosaccharides and eight amino acids participating to connect the sugar to the protein in at least 41 ways; 'this in most instances determines the nature of the carbohydrate units that will subsequently be formed by the cellular enzymatic machinery' [95].

This general information is a strong argument for far-reaching significance of the glycan part, beyond non-specific factors, such as a solubility, increase conveyed by the sugars [96]. Not long ago, 'on the whole, the role of many glycans remains obscure and the subject of speculation or controversy. 'As pointed out by Albert Neuberger in 1974: 'we are now faced with the major problem of the biological function of the glycoproteins.' However, we can be optimistic about the future … we have entered the golden age of glycoconjugates' [97]. In view of the sugars' capacity to build signals the statement that 'the significance of the glycosyl residues is to impart a discrete recognitional role on the protein' [14] proved to be prophetic.

A simple manipulation of the glycans of the serum glycoprotein ceruloplasmin as test model, i.e. removal of sialic acids to expose Gal residues, provided convincing support for this concept. In the course of studying hepatic clearance of this carrier of copper ions from serum, it was revealed 'that, in contradistinction to homologous, native ceruloplasmin, which survives for days in the serum of rabbits, intravenously injected asialoceruloplasmin [desialylation by neuraminidase of *Vibrio cholerae*] disappears from the circulation within minutes and accumulates simultaneously in the parenchymal cells of the liver' [98]. The hypothesis that the glycan is the molecular address for cell-type specific delivery was verified by covalently coupling glycopeptides from the glycoprotein fetuin to lysozyme and albumin using toluene-2,4-diisocyanate and testing clearance from serum: indeed, the glycan chains were the postal code for delivery of the synthetic neoglycoproteins to the liver, in full agreement with the observations of impact of sialic acid removal from a glycoprotein. Thus, 'the coupling technique offers a feasible method for directing the hepatic uptake of various proteins that otherwise would not 'home' to the liver'[99].

As predicted by the concept for glycan–lectin recognition, a (C-type) lectin was assumed to be expressed as a receptor for the molecular (glycan-encoded address: it was purified from detergent extracts of acetone powder prepared from the rabbit liver [100]. Homing by this type of interaction also works at the cellular-level raising monoclonal antibodies against cell preparations was crucial to unveil the first evidence for the status of endogenous lectins as cell adhesion molecules [43]. The combination of using monoclonal antibodies (against what turned out to be members of the selectin group of C-type lectins, e.g. L-selectin (CD62L, initially called MEL-14 antigen) [101,102]) and enzymatic glycan digestion on the endothelial cell surface, e.g. desialylation

with bacterial neuraminidase [103], was the breakthrough to defining the biochemical basis for the initial contact between leukocytes and vascular endothelium in lymphocyte homing (the biochemical properties that underlie the rapid and specific contact building in the bloodstream will be presented below). These two examples provide a strong incentive to present information on glycan structures associated with proteins with an eye on their potential for recognition, starting with *N*-glycans.

The conjugation of GlcNAc to the amide group of an Asn moiety in an acceptor sequon of a protein to form an *N*-glycoside explains the origin of the term *N*-glycosylation. This process is co-translational at the signal-peptide-dependent entry of the protein into the endoplasmic reticulum (ER) and involves the transfer of a preformed $GlcNAc_2Man_9Glc_3$ precursor from a lipid donor to the protein's sequon [17,104]. The 14-mer glycan is processed along the glycoproteins' routing to the Golgi by stepwise losses of sugar units, then reaching its mature form in the Golgi. The core heptasaccharide, a structure common to the three main types of *N*-glycans, is shown in the center of Figure 9. Looking at processing steps toward this core, the immediate trimming of the 14-mer glycan ensures irreversibility for the transfer reaction and thereafter the transient presence of $Glc_1Man_9GlcNAc_2$ as a ligand for the ER-based molecular lectin chaperones (calnexin and calreticulin) that have the β-sandwich fold (Figure 5) [105–107]. In the next phase, following the complete removal of Glc moieties from *N*-glycans (please note that presence of Glc in the blood precludes firm contacts between Glc as self-epitope and all other tissue lectins, explaining this strict negative selection for the composition of mature glycans), the glycoproteins are transported to the Golgi involving lectin shuttles and further trimming [108–110]. At this stage, quality control by lectin recognition also involves sensing of any accessibility of the $GlcNAc_2$ (stem) region. It is a signal for the presence of unfolded regions. This misfolded protein hereby becomes subject to ubiquitin labeling to direct it to ER-associated degradation by the SCF ubiquitin ligase complex. It contains the F-box proteins Fbs1/2 with their β-sandwich fold (shown in Figure 5) as sensor for accessibility of the $GlcNAc_2$ stem [111]. Having hereby assisted in quality control and routing mechanisms in the first phase of the passage from ER entry to secretion, the 14-mer glycan precursor has fulfilled its mission. What happens along the second part of this pathway and why?

Phenomenologically, a wide range of diversity is generated from the platform shown in Figure 9 by various means given in Figure 10: (i) core substitutions by introducing an α1,6-linked Fuc to the proximal GlcNAc of the stem or a β1,4-linked GlcNAc to the central Man unit of the *N*-glycan hepta- or pentasaccharide core that both are not starting points for elongation, (ii) introduction of branches to the core pentasaccharide by GnTs with specificity to the type of Man residue and the linkage positions (β1,2; β1,4; β1,6) up to penta-antennary complex-type *N*-glycans (with β1,6-branches accessing a larger conformational space than β1,2/4-branches due to the presence of the three-bond system (Φ, Ψ, ω-angles) with the exocyclic C5–C6 linkage of the sugar added to the glycosidic bond), (iii) addition of Galβ1,4GlcNAc (LacNAc) repeats to branches, especially the β1,6-branch and (iv) terminal tailoring to place distinct determinants at branch ends suited for recognition at different stages of their maturation, for example, yielding Gal/GalNAcβ1,4GlcNAc (LacNAc/LacdiNAc)- or α2,3/6 sialic acid/4-sulfate-terminated chains or their elaboration to histo-blood group ABH/(s)Le^x/Le^y epitopes. Each of the four routes to diversity has impacted functional pairing with lectins.

The two types of core substituents have twofold functionality: their absence/presence channel the respective substrates to a distinct pathway for either hybrid- or complex-type *N*-glycan synthesis, and they act as conformational switches in the respective *N*-glycans, hereby modulating their affinity to lectins [16,112]. The increase in the number of branches by GnT action is positively correlated to the local density of head groups. The first contact between such a cluster of glycans and a lectin is of particularly high affinity, as measured for human galectins [113]. Stepwise loading of the LacNAc branch ends of the three *N*-glycans of the glycoprotein asialofetuin with lectin up to saturation proceeds with lowered affinity [113]. With functional perspective, the ability of glycoproteins to be connected by lectin bridges to build glycoprotein-lectin aggregates depends on the extent of glycan branching [114,115] (and the type of protein architecture that underlies functional antagonism; please see below). In the case of the Glc transporter 2 that is essential for Glc-stimulated insulin secretion, reducing the number of *N*-glycan branches shortens its time of residency on the pancreatic β cell surface, assumedly by diminishing its capacity to engage in aggregate formation with a cross-linking galectin, i.e. tandem-repeat-type galectin-9 [116].

The mentioned branch extension by adding LacNAc in tandem-repeat display increases the overall length of the chain, thus the accessibility to this section of the *N*-glycan and also to the branch end. In an analogous manner, membrane-spanning lectins, such as the selectins and *s*ialic acid-binding *i*mmuno*g*lobulin-like *lec*tins (siglecs), have spacer repeats to place the terminal C-type CRD or V (=variable)-set Ig-like module as

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

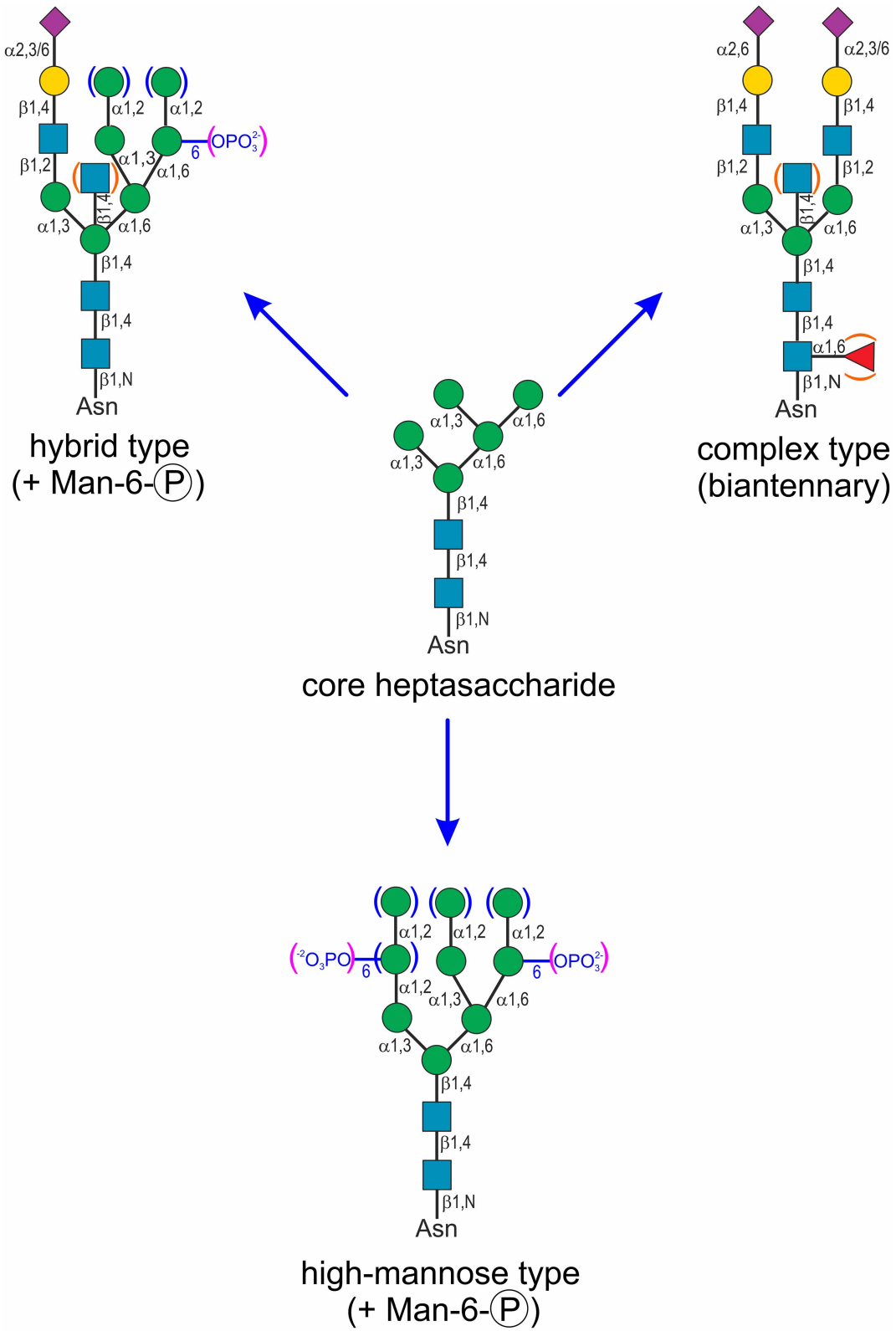


Figure 9. Structural design of vertebrate *N*-glycans. Part 1 of 2

Illustration of the three types of *N*-glycans that originate from the common core heptasaccharide. GnT-I-dependent incorporation of a GlcNAc residue into the α1,3-branch opens two routes, i.e. to the hybrid-type *N*-glycans (left, without

Figure 9. Structural design of vertebrate *N*-glycans. Part 2 of 2

trimming of the core by removing two Man moieties from the α1,6-branch) and to the complex-type *N*-glycans after the trimming and addition of the second GlcNAc residue by GnT-II into the α1,6-branch (right). Extent of branching can be increased to up to five antennae. Addition of the bisecting GlcNAc by GnT-III is a signal for pathway commitment to certain products: its presence prevents actions of α-mannosidase II (therefore directing *N*-glycan production to the hybrid-type structure), GnT-II, GnT-IV and GnT-V (determining the extent of branching) and also affects the *N*-glycan's conformational space, as the addition of the α1,6-linked fucose moiety to the core also does (core substitutions symbolized in parentheses). Together with the heptasaccharide core, its Man_6–Man_9-containing precursors constitute the class of high-mannose-type glycans that can be trimmed to paucimannosidic structures or — in yeast — extended by hypermannosylation (bottom). The Man_9-containing precursor is the platform for introducing phosphate (at the 6-position) in an enzymatic two-step process as a label for *N*-glycan recognition by the two P-type lectins that are specific for Man-6-phosphate presented in this structural context (Figure 3a). The positions of the four additional Man residues and the two phosphate groups are given in brackets.

CRD well above the bulk surface glycoproteins (16 C2-set immunoglobulin (Ig)-like repeats in the case of siglec-1 (sialoadhesin)), all to ensure perfect outreach. Not surprisingly, the relatively long branches with LacNAc repeats often participate in glycoprotein cross-linking by lectins leading to lattice formation. In addition, these repeats themselves are more than spacers. They are high-affinity docking sites for several galectins, e.g. galectin-1 at the terminal site, galectin-3 in the internal region [79,117]. This fine-tuning between galectins makes selective masking of the internal sites for galectin-3 binding possible by local glycan clustering, while maintaining the binding activity for galectin-1 [118]. Spatial accessibility and ligand properties always converge to enable avidity for lectin binding. That fine-tuning of a glycan for lectin binding is a common principle becomes most apparent by examining ligand properties along the route of *N*-glycan chain elongation. We start this survey by dealing with *N*-glycans that present GlcNAc in the context of the heptasaccharide core.

Terminal GlcNAc is recognized by several tissue lectins, a manifestation of the fundamental principle given above: the α_M-chain of the $\alpha_M\beta_2$-integrin (CR3, CD11b, Mac-1), a signal for platelet clearance by phagocytic cells [119], the C-type lectin langerin [85], the *l*iver and lymph node *s*inusoidal *e*ndothelial *c*ell C-type lec*tin* LSECtin (CLEC4G) [120], the human DCIR (CLEC4A, CLECSF6) [121] and the *d*endritic *c*ell-*s*pecific *I*CAM3-grabbing *n*onintegrin-*r*elated protein DC-SIGNR (CD299, LSIGN, CLEC4M) [122], also collectins and ficolins when presented in clusters by multi-antennary *N*-glycans [123]. Elongation of the GlcNAc-terminated branch by galactosyltransferase-mediated Gal incorporation switches off these recognition processes, making the glycan 'attractive' for new tissue receptors. Among them are the previously mentioned hepatic asialoglycoprotein receptor and the blood dendritic cell antigen 2 (BDCA-2, CD303, CLEC4C). This receptor of biantennary complex-type *N*-glycans down-regulates type I interferon production, a possible link between a glycan signal (Gal on *N*-glycans of immunoglobulins) and activity of the immune system [124]. A special role to supervise correct ligand selection is played by the Ca^{2+} in these C-type lectins [125]. Coordination bonding with the C-type lectins' Ca^{2+} involves the axial 4-OH position and also the 3-position, hereby reliably excluding Man/Glc with its equatorial 4-position. The same goal is achieved by hydrogen binding to the 4,6-positions by the galectins, teamed up with C-H/π-interactions of the B face C-H groups with π-electrons of the suitably positioned Trp residue, which explains its occurrence in the galectins' signature sequence, as already mentioned above. Alternatively, the introduction of GalNAc to form a GalNAcβ1,4GlcNAc (LacdiNAc) terminus, a characteristic compound of parasite (schistosome) surfaces shown in Figure 10, restricts galectin binding to galectin-3 [126] and proto-type galectins-2 and -7 [127]. By the way, the chemical bottom-up surface programming of nanoparticles (glycodendrimersomes) affords a remarkably tunable test system [128,129]. This protein-type-specific 4-substitution of LacdiNAc by *O*-sulfation then drastically alters the character of the glycan. Now the headgroup is a routing signal for pituitary glycoprotein hormones leading to glycoprotein uptake by the hepatic reticuloendothelial cell receptor, contact made by its β-trefoil CRD [130,131].

The sialylation of LacNAc, either in α2,3- or in α2,6-linkage, is more common than adding GalNAc. Since one of these acceptor hydroxyl groups is now blocked, contact to either the mentioned C-type lectins (in 3-position) or the galectins (in 6-position) will no longer be possible. Thus, α2,6-sialylation of LacNAc (or LacdiNAc) is a means to mask counterreceptors (for galectins), and this is used in growth regulation, for which we provide a graphic example. Deficiency in the tumor suppressor $p16^{INK4a}$ is often encountered in

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

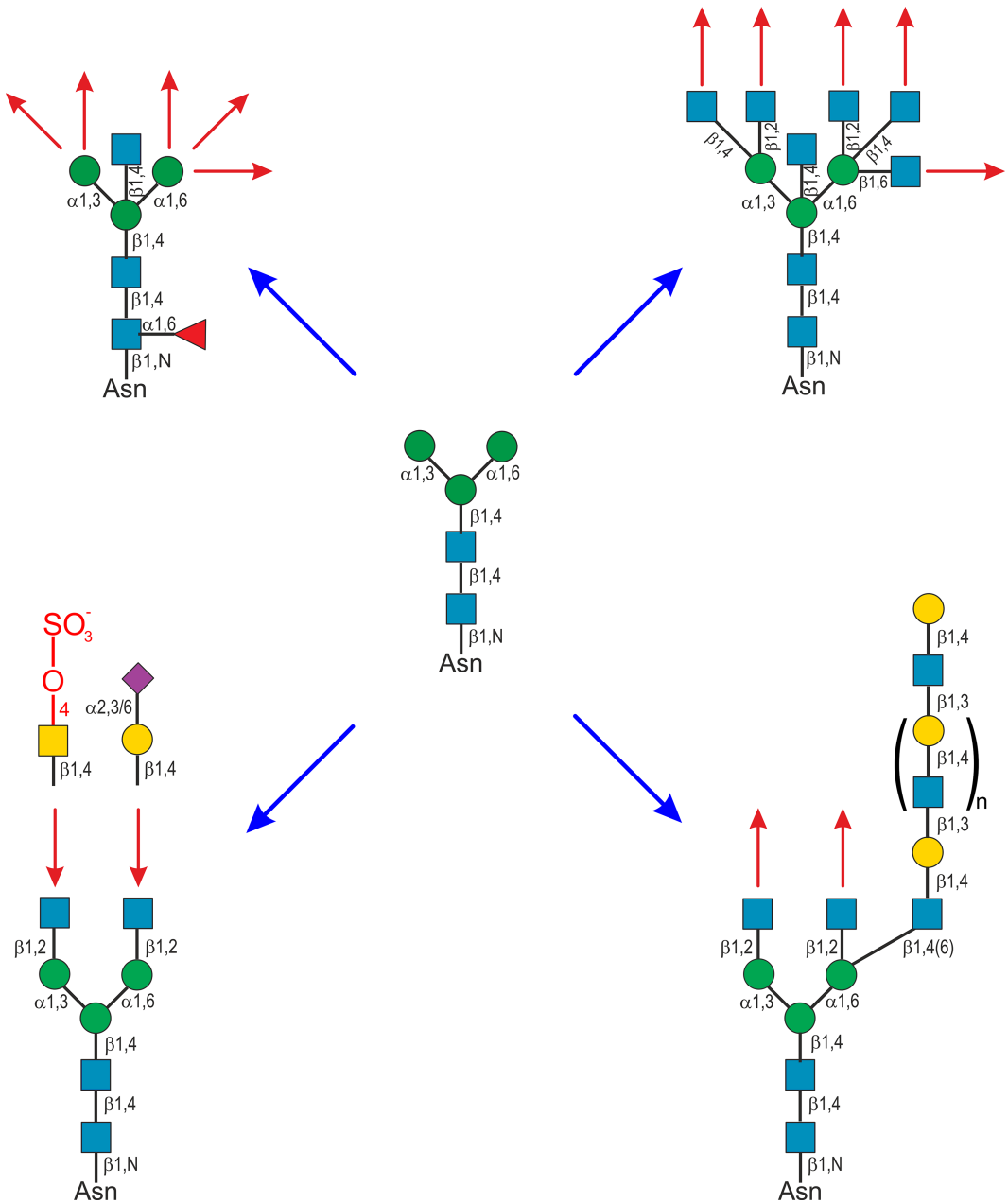


Figure 10. Examples for structural diversity in complex-type *N*-glycans.

Illustration of four routes of introducing diversity into the platform of the common core of complex-type *N*-glycans (center): it can be substituted by the bisecting GlcNAc and the core fucose units (please see also Figure 9), it can receive up to five branches by the action of specific GnTs, β1,4/6-branches in the α1,6-arm can be elongated by LacNAc repeats and branch termini can be tailored to present specific signals such as the α2,3/6-sialylated and (4-*O*-sulfated) LacdiNAc epitopes. Of course, combinations of these enzymatic reactions are possible, opening the way to a wide diversity by permutations.

pancreatic carcinoma. Among other losses and thus deviation from homeostasis, this deficiency is the cause for lifting the strict control of sialic acid biosynthesis and thus allowing high-level α2,6-sialylation. As a consequence, the $\alpha_5\beta_1$-integrin (fibronectin receptor) of pancreatic carcinoma (Capan-1) cells can no longer interact with the homodimeric (cross-linking) galectin-1. Restoration of the tumor suppressor's activity down-regulates sialic acid availability and hereby the extent of sialylation, flanked by concomitant up-regulation of galectin-1

and $\alpha_5\beta_1$-integrin presence to optimize the final effect, so that lattice formation between the integrin and the galectin induces anoikis via downstream stimulation of focal adhesion kinase and then of caspase-8 [132]. In this case, sialylation precludes glycoprotein clustering by lectin contact and bridging.

In contrast, as a docking site itself, the sialic acid is a key determinant for recognition by siglecs, siglec-2 (CD22) being a receptor for α2,6-sialylated *N*-glycans [133]. In viral (influenza) infections caused by hemagglutinin binding, the linkage positions even govern the host specificity [134]. It is, therefore, not surprising that α2,3/6-sialyltransferases show a certain degree of branch specificity in tri- and tetra-antennary *N*-glycans [135]. Similarly, the addition of another anionic substituent, i.e. the phosphate group at the 6-position in two distinct Man moieties of a high-mannose-type *N*-glycan (Figures 3 and 9), establishes a docking site for lectins, i.e. the two P-type lectins that transport hydrolytic enzymes to their sites of action in lysosomes [136,137]. What these examples are telling us, in general, is that the assembly/processing of *N*-glycans and the presence of a panel of tissue lectins facilitate multiple interactions in a specific manner. This fundamental lesson encourages us to put this idea to the test for a further class of frequent protein glycosylation already mentioned above, i.e. mucin-type *O*-glycosylation initiated by GalNAc transfer to the hydroxyl group of serine or threonine.

Like *N*-glycosylation, this type of glycan conjugation, started by 20 isoenzymes as noted above, and the elaboration of the growing glycan can take different routes in order to generate a large product panel. Common themes are stepwise additions of a sugar, occurrence of branching, addition of LacNAc repeats in the core 2(/4) branches and tailoring of branch ends to yield ABH/(s)Le^x/Le^y epitopes (Figure 11). The first product of this type of glycosylation is the α-GalNAc adduct, known as Tn antigen (CD175) [43]. The *m*acrophage *g*alactose (-binding C)-type *l*ectin (MGL, CD301, CLEC10A), closely related to the hepatic asialoglycoprotein receptor, binds this epitope [138,139]. Due to its contacts to the 3,4-hydroxyl groups by coordination bonding further processing by α2,6-sialylation (mostly by ST6GalNAc-I, one of the six α2,6-sialyltransferases using GalNAc in *O*-glycans and in gangliosides as substrate [19]) to yield sialyl (s)Tn does not impair this binding [140]. As a new contact point, the presence of the sialic acid makes siglec binding possible, as noted above for α2,6-sialylation of *N*-glycans and siglec-2, in this case first documented for siglecs-2, -3, -5 and -6 [141]. As a consequence of this binding to siglec-15, outside-in signaling is triggered so that TGF-β_1 secretion will be enhanced by the contact between monocytes and lung carcinoma cells, a factor to shape a protumoral microenvironment [142,143]. Of note, such case studies illustrate that different types of endogenous lectins take the routes to functionality as compiled in Figure 8.

Similarly acting on the levels of immune mediators, and this in positive or negative ways, paired receptors of the immunoglobulin superfamily with their siglec-like V-set extracellular CRD and intracellular pairing potential to SHP1/2 or DAP12 (in their cytoplasmic tail) also target sTn [144–146]. This disaccharide can, therefore, be considered as a signal for immunoregulation by lectin recognition. At this stage, in addition to the presence of the epitope, its density comes into play, making contacts at distinct sites likely, hereby underlining the functional significance of having an elaborate enzymatic machinery at hand [147].

This type of modification is not only present on the Tn epitope. In fact, α2,6-sialylation (by ST6GalNAc-II) operates on an alternative product of enzymatic Tn processing, i.e. β1,3-galactosylation to generate the TF disaccharide (CD176; detected by O. *T*homsen and his assistant V. *F*riedenreich, the origin of the acronym 'TF') (Figure 11). This glycan's interaction with a galectin, i.e. galectin-3 [148], appears to be relevant for breast cancer metastasis so that abrogating its ligand properties by α2,6-sialylation has been postulated to justify referring to the enzyme ST6GalNAc-II as tumor suppressor [149]. In addition to endogenous lectins, bacterial adhesins are known to be receptors for this epitope (and for other types of glycan determinants), likely contributing to guide colonization of organs such as the gut [150]. When this core 1-producing galactosyltransferase co-operates with α2,3-sialyltransferase to produce the α2,3-sialylated TF antigen and a receptor (galectin-4) is present, the resulting constellation has been implicated in the progression of castration-resistant prostate cancer [151]. The α2,3-sialylation of the TF determinant is not only conducive for Gal-4 binding. It also turns the disaccharide into a ligand for siglecs starting with siglec-1 [152]. Moreover, siglecs-3, -7, -9 and -10 share this property, the binding of the mucin MUC1 presenting this determinant on tumor cells by siglec-9 of myeloid cells favoring tumor progression by inducing the tumor-associated phenotype of macrophages with its protumoral secretion pattern [153]. At this stage, the α-GalNAc core unit can also be sialylated. Disialylation of the TF antigen establishes a potent signal for contact suited for siglec-4 (myelin-associated glycoprotein, MAG) binding, with the potential to attenuate its negative impact on axon regeneration when used as inhibitor *in vitro* [154,155]. To summarize, as seen for *N*-glycans above, each component of this synthetic route shown in Figure 11 has its own ligand properties, and there is more.

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

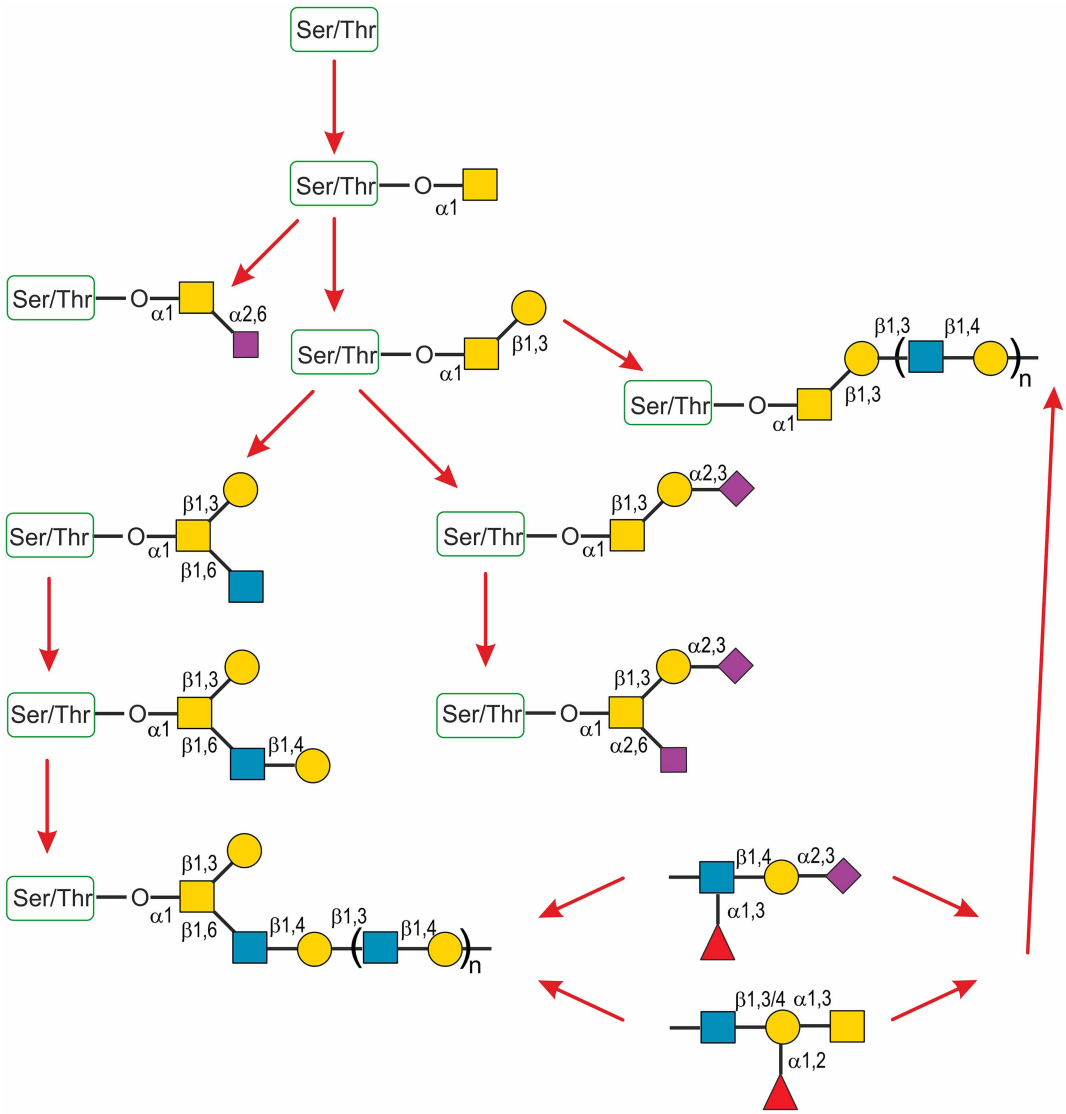


Figure 11. Structural design of vertebrate mucin-type core 1/2 *O*-glycans.

Illustration of pathways of mucin-type *O*-glycan synthesis from initial GalNAc conjugation to Ser/Thr moieties yielding core 1/2 structures.

Instead of the sialylation steps, TF antigen can alternatively become a substrate for chain elongation by LacNAc repeats or for branching to yield the core 2 structure, physiologically a double-edged sword. As seen in the *N*-glycans in an analogous manner (Figure 9), one of the three GnTs with β1,6-specificity (that recruit the exocyclic hydroxymethyl group for building a three-bond system with high-degree flexibility as in α2,6-sialylation) initiates core 2 *O*-glycan branching (Figure 11). Its significance by way of galectin binding is exemplified by the following cases on tumor growth regulation. Likely unfavorable for a patient, galectin-3 binding to core 2 *O*-glycans with LacNAc repeats on a target of a natural killer (NK) cell-activating receptor, i.e. MHC class I-related chain A, turned out to help bladder cells avoid NK cell-mediated death [156]. Galectin-1 binding to core 2 *O*-glycans, in contrast, induces prostate cancer (LNCaP) cell apoptosis [157]. Beyond the presence of branching and the introduction of LacNAc repeats, the tailoring at branch ends of core 2 *O*-glycans is significant, a recurring theme in glycan synthesis. Elaboration toward (sialyl) Lewis^x epitopes, mentioned above among the SSEAs, allows docking sites for L-selectin, here with 6-sulfation in the GlcNAc moiety [158], and for P-selectin to become prominently exposed on the branch ends of core 2 *O*-glycans [159,160]. In the context of P-selectin glycoprotein

ligand 1, co-ordinated binding to this epitope attached to Thr16 and to the two sulfated Tyr moieties (Tys7/Tys10) yields a high k_{on}-value with a K_D of 0.3 μM to enable rapid association (like anchoring a ship) to slow down cell transport in the bloodstream to a rolling [161]. The essential role of the sugar in selectin binding is underscored by the clinical consequence of impaired fucosylation, i.e. the development of leukocyte adhesion deficiency II, also called congenital disorder of glycosylation IIc [162–164]. The inclusion of a peptide portion in the recognition process, as also seen for the paired siglec-like immunoreceptor PILRα, the C-type lectin CLEC-2 or galectin-9 [165,166], is a means to gain target specificity for the functional pairing, to select, for example, CD99, podoplanin or IgE as counterreceptor in the mentioned cases.

In general, it emerges as a basic rule that synthetic pathways for glycans of glycoproteins prepare sets of signals. At each stage along the pathway, every glycan has its distinct meaning 'read' by the corresponding receptors. Having seen sialic acids to occur frequently as a building block, it is worth noting that their *O*-acetylation, 9-*O*-acetylation revealed to be essential for murine development from the two-cell stage onwards [167], and de-*O*-acetylation by *O*-acetylesterases have the potential to modulate ligand properties to selectins and siglecs [168,169], as (*N*-)acetylation affects histone properties in gene regulation with its own system of 'writers', 'readers' and 'erasers'. After all, each product has its own specific set of interaction partner(s) in our context on the side of tissue lectins. Realization of this fundamental principle also shows up in other, less frequent types of protein *O*-glycosylation. *O*-Fucosylation of epidermal growth factor- or thrombospondin type 1-like repeats integrates a site for intermolecular contact building and also intramolecular domain stabilization into the proteins [170]. In crystals, O-fucose glycans are remarkably visible so that they earned a reputation as 'surrogate amino acids' [171,172].

Another type of protein glycosylation is *O*-mannosylation. Association of multiple forms of congenital muscular dystrophy to defects in *O*-mannosylation is a definitive sign for their physiological relevance [164,173], as participation of such glycan chains in compensatory glycan restructuring in KO mice mentioned and indicated above. This type of glycosylation encompasses three core structures. Besides producing mucin-type *O*-glycan core 2-like glycans [174,175], the particular core 3 preparation with mannose phosphorylation and ribitol-5-phosphate incorporation builds the acceptor for GlcAβ1,4Xyl repeat synthesis to obtain a matriglycan on α-dystroglycan, the ligand for the Ca^{2+}-dependent contact to the β-sandwich-type laminin G-like CRD in the laminin α2 chain and in other large multidomain glycoproteins such as agrin, neurexin and perlecan, an essential factor for muscle function [176].

Structurally, the design of this polysaccharide with a disaccharide-repeat unit is similar to the respective glycosaminoglycan chains of proteoglycans [177]. As shown in Figure 1 (by gaining a shift to the skew-boat ($^2S_0$) conformer through GlcA-to-IdoA epimerization and, even more accentuated, subsequent 2-*O*-sulfation in IdoA and presence of sulfate groups in the vicinity that then facilitates a snuggly fit to the receptor) as well as in Figure 3d (by presenting the structural context of the heparin pentasaccharide with its *N*,*O*-sulfation pattern for high-affinity binding of antithrombin III), these chains, too, contribute to the pool of protein-bound glycans that become ligands. Considering the size of the toolbox, the conclusion that proteins are a versatile platform for presenting glycans as signals is indisputable. As illustrated in Figure 8, their functional pairing with lectins is significant for many aspects of cellular physiology. Thus, this surely is a part of the answer to the pertinent question 'why mammalian cell surface proteins are glycoproteins' [178], and we can apply this insight to answering the next question.

Already at the time of their discovery, (glyco)sphingolipids were understood to pose enigmas as difficult to solve as the famous mythological riddle of the Sphinx of Thebes. This explains why J. L. W. Thudichum chose to call these sugar-containing lipid compounds, which he discovered and characterized from the brain, the (glyco)sphingolipids [179,180]. The cerebrosides with their single carbohydrate head group, the neutral glycosphingolipids with glycan chains (free of a negative charge) and the gangliosides with sialic acids in their glycan chain, named in analogy to Thudichum's term cerebroside and their abundant presence in ganglia cells [181,182], indeed have been (and still are) molecules in search of functions. Providing an anchor to position lectin-binding glycans on a surface for functional pairing is emerging as one of them.

Glycosphingolipids: platform for presenting glycan-encoded information

Structurally, two characteristics of the sphingolipid backbone predestine it to become this type of general platform: (i) two lipid anchors, i.e. the C16 (palmitoyl) chain as part of sphingosine (the common precursor), and the second acyl, whose chain length can go up to C26 as source for variability, and (ii) the serine-derived hydroxyl group as an acceptor for carbohydrate in the β-anomeric position (Glc or Gal) (Figure 12). In this sense, the attached glycan chains can be formally classified as membrane-anchored *O*-glycans. Fittingly,

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

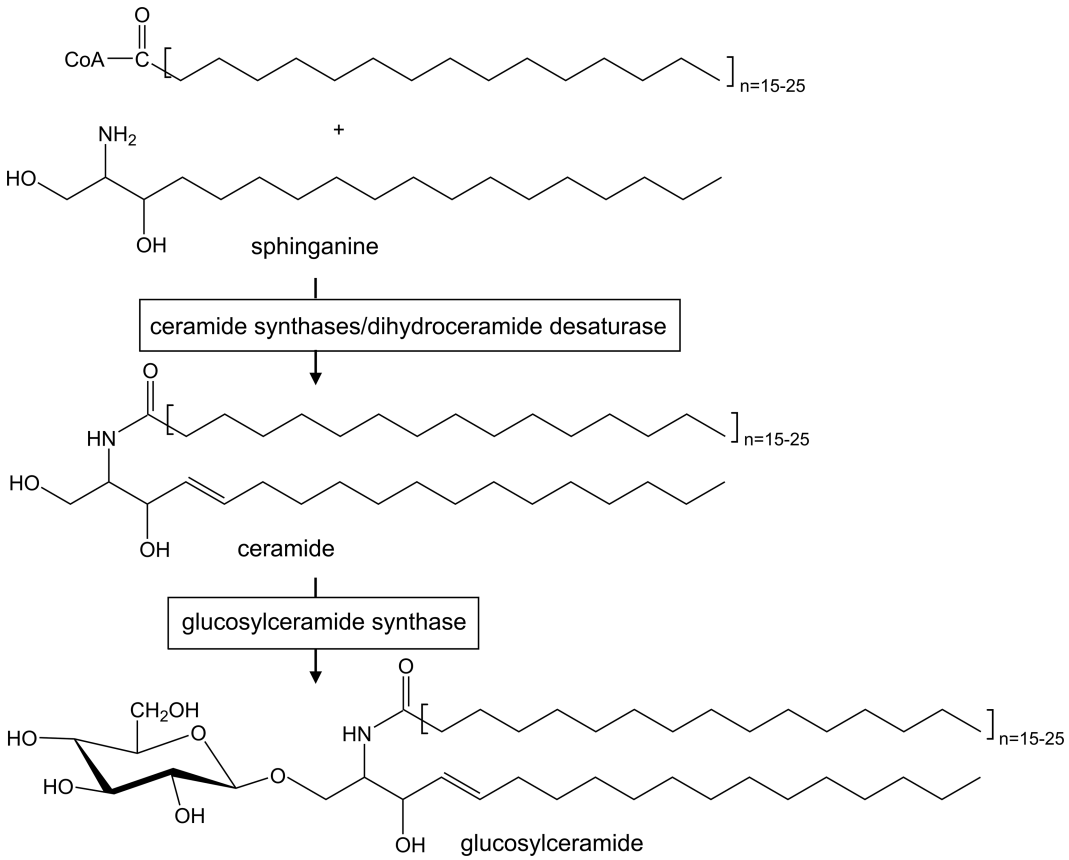


Figure 12. Structural characteristics of glycosphingolipids.

Illustration of *de novo* glycosphingolipid synthesis by adding a second acyl chain (of variable length) to sphinganine (dihydrosphingosine) by ceramide synthases (six genes in mammals), then subsequently oxidizing the dihydroceramide by a desaturase and using the serine-derived hydroxyl group of the ceramide as an acceptor site for sugar to initiate glycosylation of the ceramide. The salvage pathway recycles sphingosine directly for ceramide production by ceramide synthases.

Golgi-resident glycosyltransferases are known to be involved in acting on both on sugar chains of mucin-type glycoproteins and of glycosphingolipids. The length of the sugar chain (starting at 1 for galacto- or glucocerebrosides and the acidic (3-*O*-sulfated) galactocerebroside (sulfatide, SM4)), its sequence, the anomeric position and the linkage points of each unit in the glycan as well as the presence of sulfate/sialic acid are the determinants that assign glycosphingolipids into groups (Table 2). Since the number of known structures for both neutral and acidic glycans is in the range of 200 [183–188], the glycosphingolipids readily fulfill the criterion to offer diversity for molecular recognition processes. Clinically, cholera is a manifestation of such an interaction: the lectin part of the bacterial AB_5 toxin associates with the pentasaccharide of cell surface ganglioside GM1, and achieves this by selecting a distinct conformer (Figure 4) [189,190]. Hydrophobic contacts in the membrane among acyl chains and with proteins, together with cholesterol, are helpful to allow the glycolipids to become organized in aggregates (microdomains), with their glycan head groups in close vicinity. Adding to the branching in *N*- and *O*-glycans and organization of *O*-glycans in clusters in mucin sequence regions in glycoproteins, such an architecture microdomain is suited to reach the aim of high affinity in glycan–lectin binding.

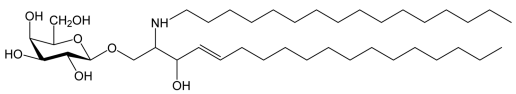
Looking first at the sulfatide, a unique cerebroside sulfotransferase catalyzes the site-specific sulfation of the Gal moiety (Gal3ST-1; [191]). *N*- and core 1 *O*-glycans of glycoproteins have their own three sulfotransferases for 3-*O*-sulfation of terminal Gal, each with its own acceptor profile [192]. They form a group within the 35 Golgi-resident human sulfotransferases using glycoconjugates as a substrate [193]. The fundamental principle of diversification is clearly at work in any section of the toolbox for glycan–lectin recognition, and the next paragraph answers the eminent question on the physiological relevance of sulfatide as docking site for lectin(s).

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

Table 2 Classification of glycosphingolipids into groups (with examples)

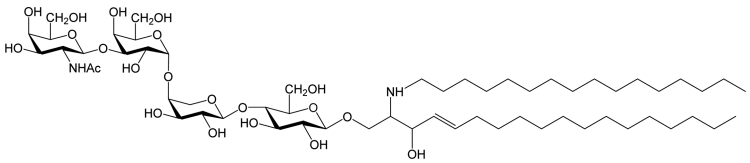
Neutral GSL

Galactosylceramide



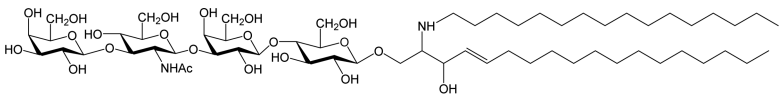
GalCer

Globotetraosylceramide



GalNAcβ1,3Galα1,4Galβ1,4Glcβ1-Cer

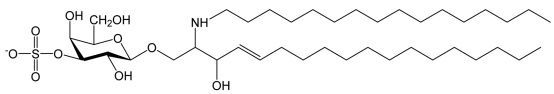
Lactotetraosylceramide



Galβ1,3GlcNAcβ1,3Galβ1,4Glcβ1-Cer

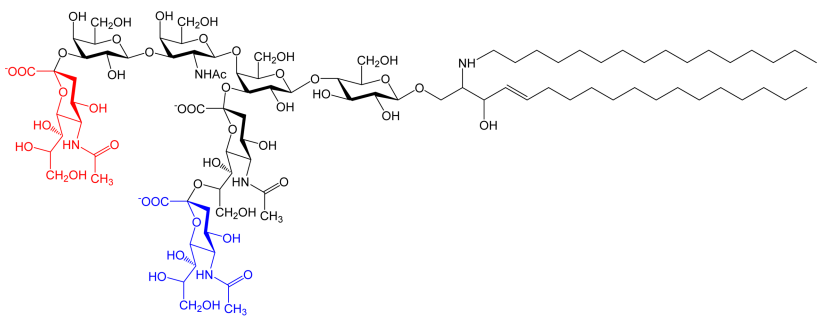
Acidic GSL

3′-Sulfogalactosylceramide (Sulfatide)



3`-Sulfo-GalCer

*Ganglioside GM1
**Ganglioside GD1a
***Ganglioside GT1b



***GM1:** Galβ1,3GalNAcβ1,4(Neu5Acα2,3)Galβ1,4GlcCer
****GD1a:** Neu5Acα2,3Galβ1,3GalNAcβ1,4(NeuAcα2,3)Galβ1,4GlcCer
*****GT1b:** Neu5Acα2,3Galβ1,3GalNAcβ1,4(Neu5Acα2,8NeuAcα2,3)Galβ1,4GlcCer

By processing galactocerebroside, a contact point for cell attachment to laminin [194,195] and for tumor cell — myeloid cell/platelet adhesion via L- and P-selectins, similar to that mediated by the sialyl Le^x epitope, is established [196–198]. Cell adhesion by binding to sulfatides is also mediated by the C-type CRD of the multimodular proteoglycans of the hyalectan/lectican group such as aggrecan or brevican [199]. The presentation of the sulfatides on the cell surface and recognition by the Ig-like receptor LMIR5 (CD300b), which teams up intracellularly with DAP12 to elicit cellular activation, is a proved route toward immunomodulation, experimentally measured in terms of induced MCP-1 production in basophils [200]. As a signal involved in the routing of distinct glycoproteins, the bivalent galectin-4, a tandem-repeat-type protein (Figure 7), brings its high-affinity ligand on detergent-resistant membranes (preferentially sulfatide species with long-chain fatty acids up to C26) together with LacNAc-bearing *N*-glycans of cargo glycoproteins. This, for example, is dipeptidylpeptidase-IV in the apical transport of polarized epithelial cells [201–204] or it is neural cell adhesion molecule L1 in its axonal sequestration, as shown in Figure 13 (Supplementary Material, Figure S1) [205].

What emerges from this inspection of interactions of the sulfatide with lectins beyond their individual significance is a functional versatility of the glycosphingolipid head group that matches the respective glycan spectrum of glycoproteins. Depending on the context and the type of receptor, the sugar signal is embedded in

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

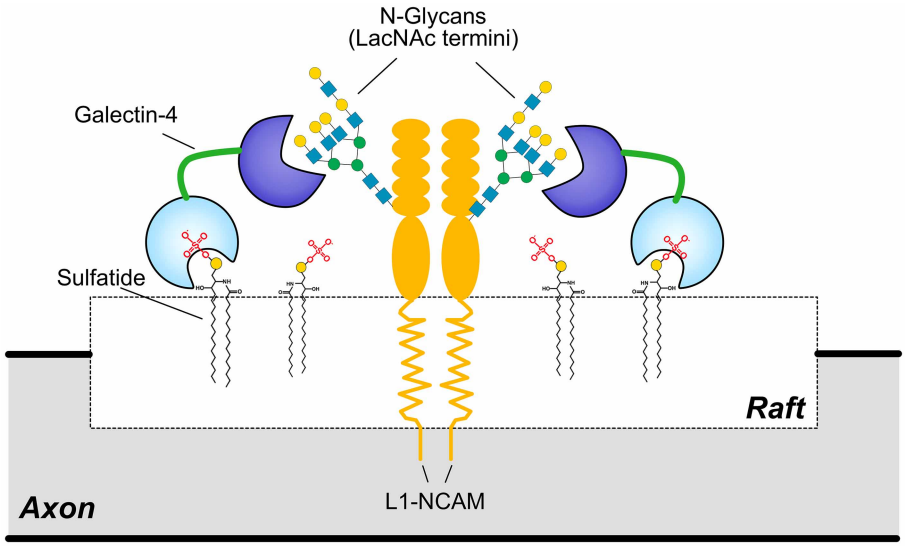


Figure 13. Galectin-4 as a bridge between sulfatide and complex-type *N*-glycans of glycoprotein(s) in cargo routing.
Illustration of the two CRDs of the tandem-repeat-type galectin-4 (please see Figure 7) in contact with their distinct counterreceptors, i.e. sulfatide (routing signal) and LacNAc-terminating complex-type *N*-glycans (cargo). As heterobifunctional cross-linker, this galectin connects its two physiological ligands to facilitate routing of cargo (glycoprotein; here L1) to particular destinations (for details, please see [205]).

pathways that reach different outcomes. This general lesson applies to glycosphingolipids with more complex carbohydrate structures than the 3-*O*-sulfated Gal in exactly the same way. Examples for such glycans are shown in Table 2. Offering sialic acid(s) as a docking site(s), the gangliosides play multifarious roles, as illustrated especially for mucin-type *O*-glycans, siglecs often acting as receptors. Siglec-7, for example, binds to di- and trisialogangliosides (and to mucin-type *O*-glycans) in the course of tumor-NK cell communication and of adhesion [206–208]. In the nervous system, siglec-4 harbors special affinity for the tri(T)- and di(D)sialylated gangliosides GT1b and GD1a, whose structures are shown in Table 2 and by using symbols in Figure 14

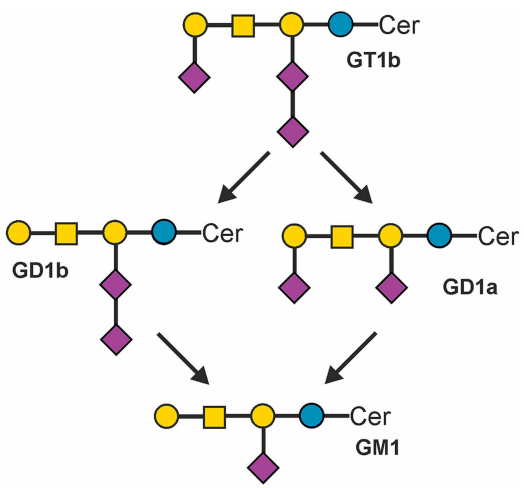


Figure 14. Dynamic enzymatic conversion of gangliosides.
Illustration how tightly controlled processing of ganglioside glycans alters the profile of binding partners for tissue lectins. Ganglioside GT1b is a substrate for cell surface ganglioside sialidase (neuraminidase 3, Neu3). Its activity yields GD1a/b. In the second Neu3-dependent step, a second sialic acid moiety is removed to generate the pentasaccharide of ganglioside GM1. In terms of recognition, a docking site for siglec-4 is turned into a counterreceptor for galectin-1.

[209,210]. This *trans*-interaction maintains regular spacing between the innermost myelin layer and the axon surfaces, and it is a channel for communication, for example, for neurite outgrowth inhibition and protective signaling to the axon [211,212]. Homodimerization (in *cis*) via the fourth and fifth Ig-like domains underlies the signaling activity in both directions [213]. Taking the scenario shown in Figure 13 one step further, *cis*-interactions between gangliosides and bi- to oligovalent lectins (such as the already mentioned cholera toxin) are possible. With broad significance, gangliosides (in clusters in microdomains) embody a platform for facilitating a switch to starting growth-regulatory programs.

At the biochemical level, the molecular nature of turning from red to green light is enzymatic desialylation, as included in Figure 14. The responsible enzyme, i.e. cell surface ganglioside sialidase (neuraminidase 3, Neu3), exhibits a remarkable degree of activity regulation. Increase in cell density was correlated to the conversion of gangliosides GT1b and GD1a to GM1, and blocking this process by a Neu3 inhibitor precluded the onset of growth control and differentiation in the neuroblastoma cells, pointing to a functional correlation [214–216]. When forcing sialic acid removal from these gangliosides by applying a bacterial sialidase as a tool, neurite outgrowth (after a Ca^{2+} influx) was induced in neuroblastoma cells, and, of note, it was prevented by the presence of an antibody against ganglioside GM1 [217].

Putting these pieces of the puzzle on the workings of this type of glycan remodeling together, the following picture emerges: a (cryptic) message (GT1b/GD1a glycans) is available on-site (that has its own meaning as siglec ligand). In the course of a cellular activation/differentiation program, the enzymatic 're-writing' (by Neu3-dependent desialylation) turns the GT1b/GD1a glycans into the GM1 pentasaccharide (Figure 14). As the mentioned case of carcinoma growth control by a tumor suppressor with glycoprotein *N*-glycans and galectin-1 has taught us, the switch between the absence/presence of sialic acid establishes a molecular switch. The new 'code word' will then be 'read' by an endogenous lectin. Regulation of its presence at the right site is a means of optimizing the response, and lectin-dependent cross-linking should be able to elicit the measured bioresponse.

The given assumptions gave research a clear direction and led to the identification of galectin-1 as ganglioside GM1 receptor that selects a glycan conformer different from that bound by cholera toxin's lectin subunits (Figure 4) [218]. Co-ordinated growth control of human neuroblastoma cells [219,220], induction of axon-like neurite outgrowth in murine neuroblastoma cells and primary cerebellar granular neurons [221] as well as attenuation of auto-immune activity of activated effector T cells by regulatory T cells [222] all include this galectin as a missing part of the puzzle along with the Neu3 activation (for review, please see [223,224]). Gangliosides GD1a/GM1 as counterreceptors, the Neu3 residing in microdomains [225] and pulling the trigger by GD1a-to-GM1 conversion as well as galectin-1 as effector form a potent team. The association of the ganglioside in microdomains to $\alpha_5\beta_1$-integrin, this glycoprotein itself being a galectin-1 counterreceptor if suitably glycosylated, then opens this route of signaling within T cell communication: autophosphorylation of focal adhesion kinase by galectin-1-dependent clustering of the $\alpha_5\beta_1$-integrin that leads to the activation of phospholipase Cγ and phosphoinositide-3 kinase and finally to Ca^{2+} influx by opening TRPC5 channels, as shown in Figure 15 (Supplementary Material, Figure S2).

Ganglioside GM1 has aspects of functionality beyond its interaction with galectins. It can also associate in *cis* with other microdomain constituents such as tropomyosin-related kinase A (TrkA), the protein tyrosine kinase receptor for nerve growth factor, thus another target of regulation by ganglioside GM1 presence (for details on the chain of events to axon growth, please see Supplementary Material, Figure S3) [226].

On the receptor site, laminin-1, too, is a binding partner for this ganglioside, an alternative starting point to trigger neurite outgrowth [227]. These examples illustrate that ganglioside GM1 has 'a multi-tasked life' and is 'a true factotum of Nature' [226]. In all these cases, the tight spatial organization of the ligand in microdomains is a factor favoring the outcome, that is positioning the right glycan (in its suited conformer; Figure 4) in the optimal spatial context in terms of local density and association to components involved in signaling. The interplay of these two factors even leads to carbohydrate–carbohydrate recognition between a ganglioside and glycan chains of a glycoprotein. The inhibitory activity of ganglioside GM3 on the epidermal growth factor receptor was attributed to an interaction between the ganglioside's trisaccharide and multiple GlcNAc termini of *N*-glycans of the glycoprotein [228]. In more general terms, this type of recognition, albeit weak, can produce sufficient strength combined with selectivity and specificity when binding partners are arranged in clusters, that is complementarily positioned [229].

In summary, when setting the insights on glycosphingolipid head group processing into relation to glycoprotein tailoring, we come up with the common rule that structural changes on glycans (a site-specific

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

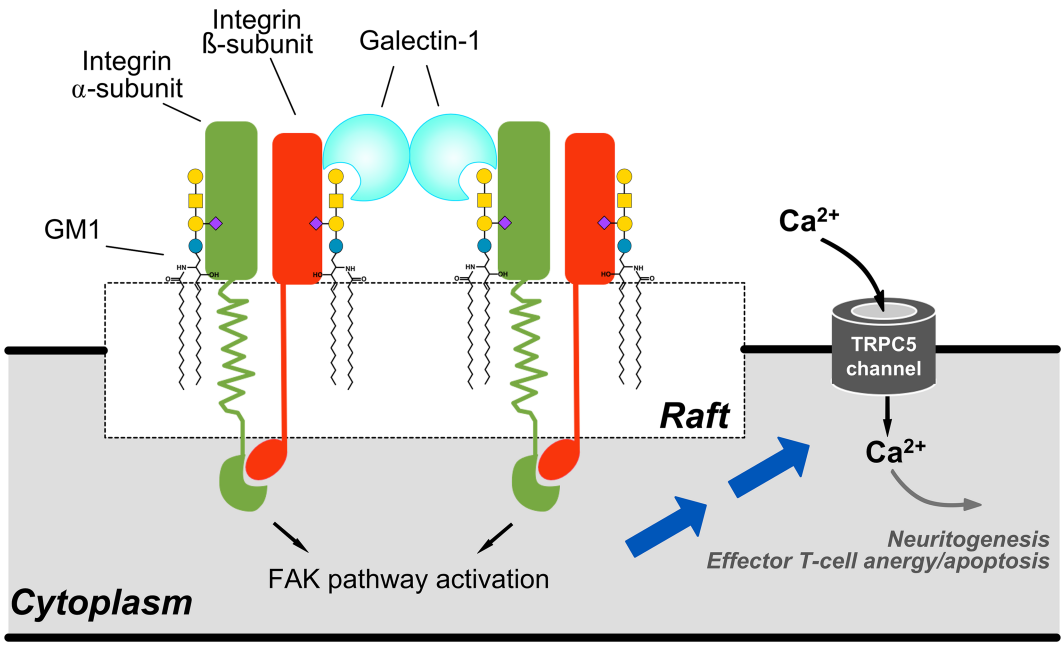


Figure 15. Functional pairing of galectin-1 and its counterreceptor(s) in inter-T cell communication/neuritogenesis.
Illustration of the chain of events from initial shift in ganglioside profile to a cellular response, starting with the enzyme Neu3. Its activity increases the presence of ganglioside GM1, as shown in Figure 14. This ganglioside associates with integrins like $\alpha_5\beta_1$-integrin and, when presented in clusters, is a high-affinity counterreceptor for galectin-1. Its cross-linking capacity forms tight aggregates with the integrin, what triggers the focal adhesion kinase pathway of signaling. It results in Ca^{2+} influx by opening TRPC5 channels (for details, please see [221,222]).

modification such as 3-*O*-sulfation of Gal or a loss of one unit from an oligomer) gain a functional dimension by lectin recognition *in situ*. Dynamic remodeling and coregulation of lectin and cognate glycan fine-tune the extent of interplay spatiotemporally. That sets of glycans of both glycoproteins and glycolipids, e.g. LacNAc termini of *N*-glycans or core M1/M2 *O*-mannose chains, core 2 *O*-glycans and ganglioside GM1, are accommodated by the same tissue lectin, here galectin-1, gives the recognition system versatility in regulation and in response profile. It would mean missing opportunities to gain optimal response patterns, if the principle to provide versatility were not operative at the level of the architecture of CRD display by lectins. This reasoning takes us from the structural description of lectins given above to an understanding of the significance of the types of their modular architecture.

Lectins: platform for multi-purpose CRD presentation

As a part of membranes, the most distal section of lectins (e.g. selectins or siglecs) certainly is a spatially privileged as the most accessible site for ligand recognition. Endocytic receptors, such as the asialoglycoprotein receptor, aggregate via their stalk region and hereby bring the CRDs into an arrangement suited for selective glycan counterreceptor binding. The molecular match between CRD distance in this lectin's trimer and the branch separation in triantennary *N*-glycans underlies the glycoside cluster effect and the logarithmic increase in affinity with linear increase in glycan density [230,231]. Linearly arrayed CRDs in tandem-repeat-type lectins, too, recognize a pattern, an effective mode of host defense. This architecture and aggregates formed via collagenous tails are well suited to trace foreign glycan signatures and to connect their detection with priming complement activation and antibody production [73–75,232,233].

Embedded into multimodular glycoproteins, a CRD can become a part of a complex interaction system within the extracellular matrix. The mentioned sulfatide binding by a C-type CRD of a proteoglycan (brevican) can be flanked by its concomitant association with hyaluronan and the ligand properties of glycosaminoglycan chains. These three adhesive elements co-operate so that the proteoglycan becomes a molecular glue between the cell surface and the matrix scaffold. As mentioned above, hyalectans or lecticans are terms for these

trimodular proteoglycans [234–236]. Modularity as a means for molecular cooperation is likewise apparent for selectins: following the rapid initial contact, the lifetime of cell–cell association under sub-critical blood flow (shear stress) is lengthened (catch bonding), likely involving an EGF-like repeat, as is, in principle, also the case for bacterial adhesins such as FimH on the distal pili position [237]. Clusters of ligands on surfaces, in adhesion or host defense, here also including soluble polysaccharides, such as β-glucans, as a platform for lectin aggregation [238], then allow patches of receptor to form, often with signaling capacity.

A *cis*-signaling-eliciting capacity is a common feature of galectins, and Figure 7 explains its structural basis. Homo- and heterobivalency and the non-covalent association via N-terminal section and/or CRD of galectin-3 enable aggregation of complexes. They can have different composition and type of topological order, hereby setting the stage for the occurrence of functional antagonism or cooperation [115,220,239,240]. Its detection gives a clue why vertebrates express galectins of the three types of architecture shown in Figure 7, and it motivates comprehensive (ga)lectin network analysis [241,242]. That the increase in galectin-1 expression as inducer of tumor cell anoikis is coupled with down-regulation of anti-apoptotic galectin-3 by the tumor suppressor $p16^{INK4a}$ [243] immediately makes sense.

Toward the same aim, that is finding answers why certain types of modular design had been included, and others excluded during phylogenesis of vertebrates, engineering to produce variant proteins of non-represented architecture, an approach termed lectinology 4.0 [244], has been performed as a test case type with CRDs of human galectins. Turning human galectin-1 into a homotetramer, a design-type found in oysters with assumed activity as pattern recognition receptor for microbial glycans (Figure 6) [245,246], the noted high-level responsiveness of cells after binding this variant already at a low density of ganglioside GM1 will severely compromise the reliability of the regulatory Neu3-GM1 axis [247]. On the other hand, strong avidity can have benefits: such an engineered protein or a galectin-3 oligomer may have potential as potent antibacterial agents in acute situations, as a galectin-3-like galectin-1 may serve as novel antagonist for medically unfavorable activities of galectin-1 [127]. Now knowing about the functional bivalency of galectin-4 in glycoprotein routing may inspire the design of new cargo transporters with the sulfatide-binding CRD giving direction to the transport, as distinct glycans, such as 4′-*O*-sulfated LacdiNAc (the equivalent of a postal code), let glycoproteins home to distinct cells in lectin-mediated delivery. Overall, shuffling of CRDs into a new structural context follows more than one incentive and also offers perspectives. This engineering approach not only provides tools to test how the architecture of CRD presentation affects activity. It also is a route that may uncover effectors with the precision of the endogenous lectins, to block undesired activities or to perform desired tasks with biomedical perspective.

Conclusions and perspectives

In principle, a molecular vocabulary with high-density coding capacity is ideal as a means for communication between cells. The special chemical properties of carbohydrates to build a large panel of linear and branched oligo- and polymeric structures from the 'letters' of the sugar alphabet, the ubiquitous occurrence of glycans in Nature and their dynamic remodeling exemplified by tightly controlled modulation of extent of sialylation on glycoproteins and glycolipids are the pivotal assets of the concept of the sugar code. That a complex machinery for glycan assembly, for modification by substitutions and for rewriting, has developed adds strong support to this fundamental concept, and these molecular messages are 'read' and 'translated' into effects by a sophisticated system of lectins (and also complementary carbohydrates). The primary structure of glycans, the local-density regulation on scaffolds by clustering or in microdomains and the shape of the glycan as well as the architectures of contact site and type of CRD presentation in lectins are factors for the specificity and the selectivity of the recognition process that determines the nature of the downstream routes of signaling. As this tutorial attests, 'evidence clearly indicates that glycans represent a largely untapped resource for biological discovery as well as unanticipated therapeutic opportunities' [2]. This statement succinctly conveys the sense of excitement and fascination of work to solve the riddles of glycans present in cellular glycoconjugates and to crack the sugar code.

Abbreviations

CRD, carbohydrate recognition domain; DCIR, dendritic cell inhibitory receptor; ER, endoplasmic reticulum; NK, natural killer; SNPs, Single nucleotide polymorphisms; SSEA, stage-specific embryonic antigens.

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

Acknowledgements

We express our sincere gratitude to the reviewers for their valuable comments that have provided inspirations to revising the original version of our manuscript.

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

References

1 Sharon, N. (1975) *Complex Carbohydrates. Their Chemistry, Biosynthesis, and Functions*, Addison-Wesley Publ. Co., Reading, MA, U.S.A.
2 Agre, P., Bertozzi, C., Bissell, M., Campbell, K.P., Cummings, R.D., Desai, U.R. et al. (2016) Training the next generation of biomedical investigators in glycosciences. *J. Clin. Invest.* **126**, 405–408 https://doi.org/10.1172/JCI85905
3 Merzendorfer, H. (2009) Chitin: structure, function and metabolism. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 217–229, Wiley-VCH, Weinheim, Germany
4 Cook, G.M.W. (1995) Glycobiology of the cell surface: the emergence of sugars as an important feature of the cell periphery. *Glycobiology* **5**, 449–461 https://doi.org/10.1093/glycob/5.5.449
5 Rademacher, T.W., Parekh, R.B. and Dwek, R.A. (1988) Glycobiology. *Annu. Rev. Biochem.* **57**, 785–838 https://doi.org/10.1146/annurev.bi.57.070188.004033
6 Gabius, H.-J. and Roth, J. (2017) An introduction to the sugar code. *Histochem. Cell Biol.* **147**, 111–117 https://doi.org/10.1007/s00418-016-1521-9
7 Reuter, G. and Gabius, H.-J. (1996) Sialic acids. Structure, analysis, metabolism, and recognition. *Biol. Chem. Hoppe Seyler* **377**, 325–342 https://doi.org/10.1515/bchm3.1996.377.6.325
8 Schauer, R. and Kamerling, J.P. (2018) Exploration of the sialic acid world. *Adv. Carbohydr. Chem. Biochem.* **75**, 1–213 https://doi.org/10.1016/bs.accb.2018.09.001
9 Roseman, S. (2001) Reflections on glycobiology. *J. Biol. Chem.* **276**, 41527–41542 https://doi.org/10.1074/jbc.R100053200
10 Peltier, P., Euzen, R., Daniellou, R., Nugier-Chauvin, C. and Ferrieres, V. (2008) Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications. *Carbohydr. Res.* **343**, 1897–1923 https://doi.org/10.1016/j.carres.2008.02.010
11 Richards, M.R. and Lowary, T.L. (2009) Chemistry and biology of galactofuranose-containing polysaccharides. *ChemBioChem* **10**, 1920–1938 https://doi.org/10.1002/cbic.200900208
12 Tefsen, B., Ram, A.F., van Die, I. and Routier, F.H. (2012) Galactofuranose in eukaryotes: aspects of biosynthesis and functional impact. *Glycobiology* **22**, 456–469 https://doi.org/10.1093/glycob/cwr144
13 Laine, R.A. (1997) The information-storing potential of the sugar code. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 1–14, Chapman & Hall, London, Weinheim
14 Winterburn, P.J. and Phelps, C.F. (1972) The significance of glycosylated proteins. *Nature* **236**, 147–151 https://doi.org/10.1038/236147a0
15 Oscarson, S. (2009) The chemist's way to synthesize glycosides. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 31–51, Wiley-VCH, Weinheim, Germany
16 Brockhausen, I. and Schachter, H. (1997) Glycosyltransferases involved in *N*- and *O*-glycan biosynthesis. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 79–113, Chapman & Hall, London, Weinheim
17 Zuber, C. and Roth, J. (2009) *N*-glycosylation. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 87–110, Wiley-VCH, Weinheim, Germany
18 Wilson, I. B. H., Paschinger, H. and Rendic, D. (2009) Glycosylation of model and 'lower' organisms. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 139–154, Wiley-VCH, Weinheim, Germany
19 Takashima, S. and Tsuji, S. (2011) Functional diversity of mammalian sialyltransferases. *Trends Glycosci. Glycotechnol.* **23**, 178–193 https://doi.org/10.4052/tigg.23.178
20 Aplin, J.D. and Jones, C.J. (2012) Fucose, placental evolution and the glycocode. *Glycobiology* **22**, 470–478 https://doi.org/10.1093/glycob/cwr156
21 Togayachi, A. and Narimatsu, H. (2012) Functional analysis of β1,3-*N*-acetylglucosaminyltransferases and regulation of immunological function by polylactosamine. *Trends Glycosci. Glycotechnol.* **24**, 95–111 https://doi.org/10.4052/tigg.24.95
22 Bhide, G.P. and Colley, K.J. (2017) Sialylation of *N*-glycans: mechanism, cellular compartmentalization and function. *Histochem. Cell Biol.* **147**, 149–174 https://doi.org/10.1007/s00418-016-1520-x
23 Corfield, A.P. (2017) Eukaryotic protein glycosylation: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* **147**, 119–147 https://doi.org/10.1007/s00418-016-1526-4
24 Roth, J. (1987) Subcellular organization of glycosylation in mammalian cells. *Biochim. Biophys. Acta* **906**, 405–436 https://doi.org/10.1016/0304-4157(87)90018-9
25 Pavelka, M. (1997) Topology of glycosylation - a histochemist's view. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 115–120, Chapman & Hall, London, Weinheim
26 Hounsell, E.F. (1997) Methods of glycoconjugate analysis. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 15–29, Chapman & Hall, London, Weinheim
27 Nakagawa, H. (2009) Analytical aspects: analysis of protein-bound glycans. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 71–83, Wiley-VCH, Weinheim, Germany
28 Alley, Jr, W.R., Mann, B.F. and Novotny, M.V. (2013) High-sensitivity analytical approaches for the structural characterization of glycoproteins. *Chem. Rev.* **113**, 2668-2732 https://doi.org/10.1021/cr3003714
29 Dam, T.K. and Brewer, C.F. (2010) Maintenance of cell surface glycan density by lectin-glycan interactions: a homeostatic and innate immune regulatory mechanism. *Glycobiology* **20**, 1061–1064 https://doi.org/10.1093/glycob/cwq084
30 Takamatsu, S., Antonopoulos, A., Ohtsubo, K., Ditto, D., Chiba, Y., Le, D.T. et al. (2010) Physiological and glycomic characterization of *N*-acetylglucosaminyltransferase-IVa and -IVb double deficient mice. *Glycobiology* **20**, 485–497 https://doi.org/10.1093/glycob/cwp200

31 Ismail, M.N., Stone, E.L., Panico, M., Lee, S.H., Luu, Y., Ramirez, K. et al. (2011) High-sensitivity *O*-glycomic analysis of mice deficient in core 2 β1,6-*N*-acetylglucosaminyltransferases. *Glycobiology* **21**, 82–98 https://doi.org/10.1093/glycob/cwq134
32 Ohtsubo, K. (2011) Biological significance of *N*-acetylglucosaminyltransferase-IV-mediated protein glycosylation on the homeostasis of cellular physiological functions. *Trends Glycosci. Glycotechnol.* **23**, 103–105 https://doi.org/10.4052/tigg.23.103
33 Mkhikian, H., Mortales, C.L., Zhou, R.W., Khachikyan, K., Wu, G., Haslam, S.M. et al. (2016) Golgi self-correction generates bioequivalent glycans to preserve cellular homeostasis. *eLife* **5**, e14814 https://doi.org/10.7554/eLife.14814
34 Bennett, E.P., Mandel, U., Clausen, H., Gerken, T.A., Fritz, T.A. and Tabak, L.A. (2012) Control of mucin-type *O*-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22**, 736–756 https://doi.org/10.1093/glycob/cwr182
35 Raman, J., Guan, Y., Perrine, C.L., Gerken, T.A. and Tabak, L.A. (2012) UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferases: completion of the family tree. *Glycobiology* **22**, 768–777 https://doi.org/10.1093/glycob/cwr183
36 de Las Rivas, M., Lira-Navarrete, E., Gerken, T.A. and Hurtado-Guerrero, R. (2019) Polypeptide GalNAc-Ts: from redundancy to specificity. *Curr. Opin. Struct. Biol.* **56**, 87–96 https://doi.org/10.1016/j.sbi.2018.12.007
37 Brockhausen, I., Schutzbach, J. and Kuhns, W. (1998) Glycoproteins and their relationship to human disease. *Acta Anat.* **161**, 36–78 https://doi.org/10.1159/000046450
38 Reuter, G. and Gabius, H.-J. (1999) Eukaryotic glycosylation: whim of nature or multipurpose tool? *Cell. Mol. Life Sci.* **55**, 368–422 https://doi.org/10.1007/s000180050298
39 Corfield, A.P. (2015) Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta* **1850**, 236–252 https://doi.org/10.1016/j.bbagen.2014.05.003
40 Solter, D. and Knowles, B.B. (1978) Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *Proc. Natl Acad. Sci. U.S.A.* **75**, 5565–5569 https://doi.org/10.1073/pnas.75.11.5565
41 Feizi, T. (1985) Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens. *Nature* **314**, 53–57 https://doi.org/10.1038/314053a0
42 Fenderson, B.A., Eddy, E.M. and Hakomori, S.-i. (1990) Glycoconjugate expression during embryogenesis and its biological significance. *BioEssays* **12**, 173–179 https://doi.org/10.1002/bies.950120406
43 Gabius, H.-J., Kaltner, H., Kopitz, J. and André, S. (2015) The glycobiology of the CD system: a dictionary for translating marker designations into glycan/lectin structure and function. *Trends Biochem. Sci.* **40**, 360–376 https://doi.org/10.1016/j.tibs.2015.03.013
44 Carver, J.P. (1993) Oligosaccharides: how can flexible molecules act as signals? *Pure Appl. Chem.* **65**, 763–770 https://doi.org/10.1351/pac199365040763
45 Gabius, H.-J. (1998) The how and why of protein-carbohydrate interaction: a primer to the theoretical concept and a guide to application in drug design. *Pharmaceut. Res.* **15**, 23–30 https://doi.org/10.1023/A:1011936300845
46 von der Lieth, C.-W., Siebert, H.-C., Kozár, T., Burchert, M., Frank, M., Gilleron, M. et al. (1998) Lectin ligands: new insights into their conformations and their dynamic behavior and the discovery of conformer selection by lectins. *Acta Anat.* **161**, 91–109 https://doi.org/10.1159/000046452
47 Hardy, B.J. (1997) The glycosidic linkage flexibility and time-scale similarity hypotheses. *J. Mol. Struct.* **395–396**, 187–200 https://doi.org/10.1016/S0166-1280(96)04866-X
48 Quiocho, F.A. (1986) Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu. Rev. Biochem.* **55**, 287–315 https://doi.org/10.1146/annurev.bi.55.070186.001443
49 Rini, J.M. (1995) Lectin structure. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577 https://doi.org/10.1146/annurev.bb.24.060195.003003
50 Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. and Solís, D. (2011) From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.* **36**, 298–313 https://doi.org/10.1016/j.tibs.2011.01.005
51 Solís, D., Bovin, N.V., Davis, A.P., Jiménez-Barbero, J., Romero, A., Roy, R. et al. (2015) A guide into glycosciences: how chemistry, biochemistry and biology cooperate to crack the sugar code. *Biochim. Biophys. Acta* **1850**, 186–235 https://doi.org/10.1016/j.bbagen.2014.03.016
52 Barondes, S.H. (1997) Galectins: a personal overview. *Trends Glycosci. Glycotechnol.* **9**, 1–7 https://doi.org/10.4052/tigg.9.1
53 Neuberger, A. (1995) Preface. In *Glycoproteins* (Montreuil, J., Vliegenthart, J.F.G. and Schachter, H., eds), pp. v–vii, Elsevier, Amsterdam
54 Gabius, H.-J. (ed.) (2009) *The Sugar Code. Fundamentals of Glycosciences*, Wiley-VCH, Weinheim, Germany
55 Hart, G.W. (2013) Thematic minireview series on glycobiology and extracellular matrices: glycan functions pervade biology at all levels. *J. Biol. Chem.* **288**, 6903 https://doi.org/10.1074/jbc.R113.453977
56 Boyd, W.C. and Shapleigh, E. (1954) Specific precipitating activity of plant agglutinins (lectins). *Science* **119**, 419 https://doi.org/10.1126/science.119.3091.419
57 Kocourek, J. (1986) Historical background. In *The Lectins. Properties, Functions and Applications in Biology and Medicine* (Liener, I.E., Sharon, N. and Goldstein, I.J., eds), pp. 1–32, Academic Press, New York
58 Kilpatrick, D.C. and Green, C. (1992) Lectins as blood typing reagents. *Adv. Lectin Res.* **5**, 51–94
59 Barondes, S.H. (1988) Bifunctional properties of lectins: lectins redefined. *Trends Biochem. Sci.* **13**, 480–482 https://doi.org/10.1016/0968-0004(88)90235-6
60 Watkins, W.M. (1999) A half century of blood-group antigen research: some personal recollections. *Trends Glycosci. Glycotechnol.* **11**, 391–411 https://doi.org/10.4052/tigg.11.391
61 Kaltner, H., García Caballero, G., Ludwig, A.-K., Manning, J.C. and Gabius, H.-J. (2018) From glycophenotyping by (plant) lectin histochemistry to defining functionality of glycans by pairing with endogenous lectins. *Histochem. Cell Biol.* **149**, 547–568 https://doi.org/10.1007/s00418-018-1676-7
62 Spicer, S.S. and Schulte, B.A. (1992) Diversity of cell glycoconjugates shown histochemically: a perspective. *J. Histochem. Cytochem.* **40**, 1–38 https://doi.org/10.1177/40.1.1370305
63 Danguy, A., Akif, F., Pajak, B. and Gabius, H.-J. (1994) Contribution of carbohydrate histochemistry to glycobiology. *Histol. Histopathol.* **9**, 155–171 PMID:8003812
64 Roth, J. (2011) Lectins for histochemical demonstration of glycans. *Histochem. Cell Biol.* **136**, 117–130 https://doi.org/10.1007/s00418-011-0848-5
65 Tang, H., Hsueh, P., Kletter, D., Bern, M. and Haab, B. (2015) The detection and discovery of glycan motifs in biological samples using lectins and antibodies: new methods and opportunities. *Adv. Cancer Res.* **126**, 167–202 https://doi.org/10.1016/bs.acr.2014.11.003

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

66 Manning, J.C., Romero, A., Habermann, F.A., García Caballero, G., Kaltner, H. and Gabius, H.-J. (2017) Lectins: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* **147**, 199–222 https://doi.org/10.1007/s00418-016-1524-6
67 Fujimoto, Z., Tateno, H. and Hirabayashi, J. (2014) Lectin structures: classification based on the 3-D structures. *Methods Mol. Biol.* **1200**, 579–606 https://doi.org/10.1007/978-1-4939-1292-6_46
68 Wesener, D.A., Wangkanont, K., McBride, R., Song, X., Kraft, M.B., Hodges, H.L. et al. (2015) Recognition of microbial glycans by human intelectin-1. *Nat. Struct. Mol. Biol.* **22**, 603–610 https://doi.org/10.1038/nsmb.3053
69 Lehotzky, R.E., Partch, C.L., Mukherjee, S., Cash, H.L., Goldman, W.E., Gardner, K.H. et al. (2010) Molecular basis for peptidoglycan recognition by a bactericidal lectin. *Proc. Natl Acad. Sci. U.S.A.* **107**, 7722–7727 https://doi.org/10.1073/pnas.0909449107
70 Matsunaga, I. and Moody, D.B. (2009) Mincle is a long sought receptor for mycobacterial cord factor. *J. Exp. Med.* **206**, 2865–2868 https://doi.org/10.1084/jem.20092533
71 Ishikawa, E., Ishikawa, T., Morita, Y.S., Toyonaga, K., Yamada, H., Takeuchi, O. et al. (2009) Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J. Exp. Med.* **206**, 2879–2888 https://doi.org/10.1084/jem.20091750
72 Furukawa, A., Kamishikiryo, J., Mori, D., Toyonaga, K., Okabe, Y., Toji, A. et al. (2013) Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL. *Proc. Natl Acad. Sci. U.S.A.* **110**, 17438–17443 https://doi.org/10.1073/pnas.1312649110
73 Kilpatrick, D.C. (2002) Mannan-binding lectin: clinical significance and applications. *Biochim. Biophys. Acta* **1572**, 401–413 https://doi.org/10.1016/S0304-4165(02)00321-5
74 Lu, J., Teh, C., Kishore, U. and Reid, K.B.M. (2002) Collectins and ficolins: sugar pattern recognition molecules of the mammalian innate immune system. *Biochim. Biophys. Acta* **1572**, 387–400 https://doi.org/10.1016/S0304-4165(02)00320-3
75 Martinez-Pomares, L. (2012) The mannose receptor. *J. Leukoc. Biol.* **92**, 1177–1186 https://doi.org/10.1189/jlb.0512231
76 Gilleron, M., Siebert, H.-C., Kaltner, H., von der Lieth, C.-W., Kozár, T., Halkes, K.M. et al. (1998) Conformer selection and differential restriction of ligand mobility by a plant lectin. Conformational behavior of Galβ1-3GlcNAcβ1-R, Galβ1-3GalNAcβ1-R and Galβ1-2Galβ1-R′ in the free state and complexed with mistletoe lectin as revealed by random walk and conformational clustering molecular mechanics calculations, molecular dynamics simulations and nuclear Overhauser experiments. *Eur. J. Biochem.* **252**, 416–427 https://doi.org/10.1046/j.1432-1327.1998.2520416.x
77 Asensio, J.L., Espinosa, J.F., Dietrich, H., Cañada, F.J., Schmidt, R.R., Martín-Lomas, M. et al. (1999) Bovine heart galectin-1 selects a distinct *(syn)* conformation of C-lactose, a flexible lactose analogue. *J. Am. Chem. Soc.* **121**, 8995–9000 https://doi.org/10.1021/ja990601u
78 Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M. et al. (2002) Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim. Biophys. Acta* **1572**, 232–254 https://doi.org/10.1016/S0304-4165(02)00311-2
79 Iwaki, J. and Hirabayashi, J. (2018) Carbohydrate-binding specificity of human galectins: an overview by frontal affinity chromatography. *Trends Glycosci. Glycotechnol.* **30**, SE137–SE153 https://doi.org/10.4052/tigg.1728.1SE
80 Stowell, S.R., Arthur, C.M., Mehta, P., Slanina, K.A., Blixt, O., Leffler, H. et al. (2008) Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens. *J. Biol. Chem.* **283**, 10109–10123 https://doi.org/10.1074/jbc.M709545200
81 Suzuki, N., Yamamoto, K., Toyoshima, S., Osawa, T. and Irimura, T. (1996) Molecular cloning and expression of cDNA encoding human macrophage C-type lectin. *J. Immunol.* **156**, 128–135 PMID:8598452
82 Tsuiji, M., Fujimori, M., Ohashi, Y., Higashi, N., Onami, T.M., Hedrick, S.M. et al. (2002) Molecular cloning and characterization of a novel mouse macrophage C-type lectin, mMGL2, which has a distinct carbohydrate specificity from mMGL1. *J. Biol. Chem.* **277**, 28892–28901 https://doi.org/10.1074/jbc.M203774200
83 Flornes, L.M., Bryceson, Y.T., Spurkland, A., Lorentzen, J.C., Dissen, E. and Fossum, S. (2004) Identification of lectin-like receptors expressed by antigen presenting cells and neutrophils and their mapping to a novel gene complex. *Immunogenetics* **56**, 506–517 https://doi.org/10.1007/s00251-004-0714-x
84 Garcia-Vallejo, J.J. and van Kooyk, Y. (2013) The physiological role of DC-SIGN: a tale of mice and men. *Trends Immunol.* **34**, 482–486 https://doi.org/10.1016/j.it.2013.03.001
85 Feinberg, H., Rowntree, T.J., Tan, S.L., Drickamer, K., Weis, W.I. and Taylor, M.E. (2013) Common polymorphisms in human langerin change specificity for glycan ligands. *J. Biol. Chem.* **288**, 36762–36771 https://doi.org/10.1074/jbc.M113.528000
86 Ruiz, F.M., Scholz, B.A., Buzamet, E., Kopitz, J., André, S., Menéndez, M. et al. (2014) Natural single amino acid polymorphism (F19Y) in human galectin-8: detection of structural alterations and increased growth-regulatory activity on tumor cells. *FEBS J.* **281**, 1446–1464 https://doi.org/10.1111/febs.12716
87 Kasai, K.-i. and Hirabayashi, J. (1996) Galectins: a family of animal lectins that decipher glycocodes. *J. Biol. Chem.* **119**, 1–8 https://doi.org/10.1093/oxfordjournals.jbchem.a021192
88 Gabius, H.-J. (1997) Animal lectins. *Eur. J. Biochem.* **243**, 543–576 https://doi.org/10.1111/j.1432-1033.1997.t01-1-00543.x
89 Cooper, D.N.W. (2002) Galectinomics: finding themes in complexity. *Biochim. Biophys. Acta* **1572**, 209–231 https://doi.org/10.1016/S0304-4165(02)00310-0
90 Kaltner, H., Toegel, S., García Caballero, G., Manning, J.C., Ledeen, R.W. and Gabius, H.-J. (2017) Galectins: their network and roles in immunity/tumor growth control. *Histochem. Cell Biol.* **147**, 239–256 https://doi.org/10.1007/s00418-016-1522-8
91 Gready, J.N. and Zelensky, A.N. (2009) Routes in lectin evolution: case study on the C-type lectin-like domains. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 329–346, Wiley-VCH, Weinheim, Germany
92 Gabius, H.-J., Springer, W.R. and Barondes, S.H. (1985) Receptor for the cell binding site of discoidin I. *Cell* **42**, 449–456 https://doi.org/10.1016/0092-8674(85)90102-3
93 Eichwald, E. (1865) Beiträge zur Chemie der gewebbildenden Substanzen und ihrer Abkömmlinge. I. Ueber das Mucin, besonders der Weinbergschnecke. *Ann. Chem. Pharm.* **134**, 177–211 https://doi.org/10.1002/jlac.18651340207
94 Johansen, P.G., Marshall, R.D. and Neuberger, A. (1961) Carbohydrates in protein. 3. The preparation and some of the properties of a glycopeptide from hen's-egg albumin. *Biochem. J.* **78**, 518–527 https://doi.org/10.1042/bj0780518
95 Spiro, R.G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R–56R https://doi.org/10.1093/glycob/12.4.43R
96 Sharon, N. and Lis, H. (1997) Glycoproteins: structure and function. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 133–162, Chapman & Hall, London, Weinheim

97 Montreuil, J. (1995) The history of glycoprotein research, a personal view. In *Glycoproteins* (Montreuil, J., Vliegenthart, J.F.G. and Schachter, H., eds), pp. 1–12, Elsevier, Amsterdam, The Netherlands
98 Morell, A.G., Irvine, R.A., Sternlieb, I., Scheinberg, I.H. and Ashwell, G. (1968) Physical and chemical studies on ceruloplasmin. V. Metabolic studies on sialic acid-free ceruloplasmin *in vivo*. *J. Biol. Chem.* **243**, 155–159 PMID:5635941
99 Rogers, J.C. and Kornfeld, S. (1971) Hepatic uptake of proteins coupled to fetuin glycopeptide. *Biochem. Biophys. Res. Commun.* **45**, 622–629 https://doi.org/10.1016/0006-291X(71)90462-1
100 Hudgin, R.L., Pricer, W.E.J., Ashwell, G., Stockert, R.J. and Morell, A.G. (1974) The isolation and properties of a rabbit liver binding protein specific for asialoglycoproteins. *J. Biol. Chem.* **249**, 5536–5543 PMID:4370480
101 Gallatin, W.M., Weissman, I.L. and Butcher, E.C. (1983) A cell-surface molecule involved in organ-specific homing of lymphocytes. *Nature* **304**, 30–34 https://doi.org/10.1038/304030a0
102 Kaltner, H. and Gabius, H.-J. (2019) Sensing glycans as biochemical messages by tissue lectins: the sugar code at work in vascular biology. *Thromb. Haemost.* **119**, 517–533 https://doi.org/10.1055/s-0038-1676968
103 Rosen, S.D., Singer, M.S., Yednock, T.A. and Stoolman, L.M. (1985) Involvement of sialic acid on endothelial cells in organ-specific lymphocyte recirculation. *Science* **228**, 1005–1007 https://doi.org/10.1126/science.4001928
104 Kelleher, D.J. and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47R–62R https://doi.org/10.1093/glycob/cwj066
105 Roth, J. (2002) Protein *N*-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem. Rev.* **102**, 285–303 https://doi.org/10.1021/cr000423j
106 Aebi, M., Bernasconi, R., Clerc, S. and Molinari, M. (2010) *N*-Glycan structures: recognition and processing in the ER. *Trends Biochem. Sci.* **35**, 74–82 https://doi.org/10.1016/j.tibs.2009.10.001
107 Pearse, B.R. and Hebert, D.N. (2010) Lectin chaperones help direct the maturation of glycoproteins in the endoplasmic reticulum. *Biochim. Biophys. Acta* **1803**, 684–693 https://doi.org/10.1016/j.bbamcr.2009.10.008
108 Yamashita, K., Hara-Kuge, S. and Ohkura, T. (1999) Intracellular lectins associated with *N*-linked glycoprotein traffic. *Biochim. Biophys. Acta* **1473**, 147–160 https://doi.org/10.1016/S0304-4165(99)00175-0
109 Hauri, H.P., Nufer, O., Breuza, L., Tekaya, H.B. and Liang, L. (2002) Lectins and protein traffic early in the secretory pathway. *Biochem. Soc. Symp.* **69**, 73–82 https://doi.org/10.1042/bss0690073
110 Roth, J. and Zuber, C. (2017) Quality control of glycoprotein folding and ERAD: the role of *N*-glycan handling, EDEM1 and OS-9. *Histochem. Cell Biol.* **147**, 269–284 https://doi.org/10.1007/s00418-016-1513-9
111 Yoshida, Y. (2007) F-box proteins that contain sugar-binding domains. *Biosci. Biotechnol. Biochem.* **71**, 2623–2631 https://doi.org/10.1271/bbb.70074
112 André, S., Kozár, T., Kojima, S., Unverzagt, C. and Gabius, H.-J. (2009) From structural to functional glycomics: core substitutions as molecular switches for shape and lectin affinity of *N*-glycans. *Biol. Chem.* **390**, 557–565 https://doi.org/10.1515/BC.2009.072
113 Dam, T.K., Gabius, H.-J., André, S., Kaltner, H., Lensch, M. and Brewer, C.F. (2005) Galectins bind to the multivalent glycoprotein asialofetuin with enhanced affinities and a gradient of decreasing binding constants. *Biochemistry* **44**, 12564–12571 https://doi.org/10.1021/bi051144z
114 Gupta, D., Kaltner, H., Dong, X., Gabius, H.-J. and Brewer, C.F. (1996) Comparative cross-linking activities of lactose-specific plant and animal lectins and a natural lactose-binding immunoglobulin G fraction from human serum with asialofetuin. *Glycobiology* **6**, 843–849 https://doi.org/10.1093/glycob/6.8.843
115 Ahmad, N., Gabius, H.-J., André, S., Kaltner, H., Sabesan, S., Roy, R. et al. (2004) Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *J. Biol. Chem.* **279**, 10841–10847 https://doi.org/10.1074/jbc.M312834200
116 Ohtsubo, K., Takamatsu, S., Minowa, M.T., Yoshida, A., Takeuchi, M. and Marth, J.D. (2005) Dietary and genetic control of glucose transporter 2 glycosylation promotes insulin secretion in suppressing diabetes. *Cell* **123**, 1307–1321 https://doi.org/10.1016/j.cell.2005.09.041
117 Ahmad, N., Gabius, H.-J., Kaltner, H., André, S., Kuwabara, I., Liu, F.-T. et al. (2002) Thermodynamic binding studies of cell surface carbohydrate epitopes to galectins-1, -3 and -7. Evidence for differential binding specificities. *Can. J. Chem.* **80**, 1096–1104 https://doi.org/10.1139/v02-162
118 St-Pierre, C., Manya, H., Ouellet, M., Clark, G.F., Endo, T., Tremblay, M.J. et al. (2011) Host-soluble galectin-1 promotes HIV-1 replication through a direct interaction with glycans of viral gp120 and host CD4. *J. Virol.* **85**, 11742–11751 https://doi.org/10.1128/JVI.05351-11
119 Hoffmeister, K. and Falet, H. (2009) Platelet glycoproteins as lectin in hematology. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 485–493, Wiley-VCH, Weinheim, Germany
120 Pipirou, Z., Powlesland, A.S., Steffen, I., Pohlmann, S., Taylor, M.E. and Drickamer, K. (2011) Mouse LSECtin as a model for a human Ebola virus receptor. *Glycobiology* **21**, 806–812 https://doi.org/10.1093/glycob/cwr008
121 Nagae, M., Ikeda, A., Hanashima, S., Kojima, T., Matsumoto, N., Yamamoto, K. et al. (2016) Crystal structure of human dendritic cell inhibitory receptor C-type lectin domain reveals the binding mode with *N*-glycan. *FEBS Lett.* **590**, 1280–1288 https://doi.org/10.1002/1873-3468.12162
122 Davis, C.W., Mattei, L.M., Nguyen, H.Y., Ansarah-Sobrinho, C., Doms, R.W. and Pierson, T.C. (2006) The location of asparagine-linked glycans on West Nile virions controls their interactions with CD209 (dendritic cell-specific ICAM-3 grabbing nonintegrin). *J. Biol. Chem.* **281**, 37183–37194 https://doi.org/10.1074/jbc.M605429200
123 Kawasaki, N. and Kawasaki, T. (2010) Recognition of endogenous ligands by C-type lectins: interaction of serum mannan-binding protein with tumor-associated oligosaccharide epitopes. *Trends Glycosci. Glycotechnol.* **22**, 141–151 https://doi.org/10.4052/tigg.22.141
124 Kim, J.W., Budzak, J., Liu, Y., Jegouzo, S.A.F., Drickamer, K. and Taylor, M.E. (2018) Identification of serum glycoprotein ligands for the immunomodulatory receptor blood dendritic cell antigen 2. *Glycobiology* **28**, 592–600 https://doi.org/10.1093/glycob/cwy050
125 Gabius, H.-J. (2011) The how and why of Ca^{2+} involvement in lectin activity. *Trends Glycosci. Glycotechnol.* **23**, 168–177 https://doi.org/10.4052/tigg.23.168
126 van den Berg, T.K., Honing, H., Franke, N., van Remoortere, A., Schiphorst, W.E.C.M., Liu, F.-T. et al. (2004) LacdiNAc-glycans constitute a parasite pattern for galectin-3-mediated immune recognition. *J. Immunol.* **173**, 1902–1907 https://doi.org/10.4049/jimmunol.173.3.1902
127 Ludwig, A.-K., Michalak, M., Xiao, Q., Gilles, U., Medrano, F.J., Ma, H. et al. (2019) Design-functionality relationships for adhesion/growth-regulatory galectins. *Proc. Natl Acad. Sci. U.S.A.* **116**, 2837–2842 https://doi.org/10.1073/pnas.1813515116
128 Percec, V., Leowanawat, P., Sun, H.J., Kulikov, O., Nusbaum, C.D., Tran, T.M. et al. (2013) Modular synthesis of amphiphilic Janus glycodendrimers and their self-assembly into glycodendrimersomes and other complex architectures with bioactivity to biomedically relevant lectins. *J. Am. Chem. Soc.* **135**, 9055–9077 https://doi.org/10.1021/ja403323y

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

129 Xiao, Q., Ludwig, A.-K., Romano, C., Buzzacchera, I., Sherman, S.E., Vetro, M. et al. (2018) Exploring functional pairing between surface glycoconjugates and human galectins using programmable glycodendrimersomes. *Proc. Natl Acad. Sci. U.S.A.* **115**, E2509–E2518 https://doi.org/10.1073/pnas.1720055115

130 Fiete, D., Beranek, M.C. and Baenziger, J.U. (1997) The macrophage/endothelial cell mannose receptor cDNA encodes a protein that binds oligosaccharides terminating with SO_4-4-GalNAcβ1,4GlcNAcβ or Man at independent sites. *Proc. Natl Acad. Sci. U.S.A.* **94**, 11256–11261 https://doi.org/10.1073/pnas.94.21.11256

131 Hooper, L.V., Manzella, S.M. and Baenziger, J.U. (1997) The biology of sulfated oligosaccharides. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 261–276, Chapman & Hall, London, Weinheim

132 Amano, M., Eriksson, H., Manning, J.C., Detjen, K.M., André, S., Nishimura, S.-I. et al. (2012) Tumour suppressor $p16^{INK4a}$: anoikis-favouring decrease in *N/O*-glycan/cell surface sialylation by down-regulation of enzymes in sialic acid biosynthesis in tandem in a pancreatic carcinoma model. *FEBS J.* **279**, 4062–4080 https://doi.org/10.1111/febs.12001

133 Angata, T. and Brinkman-van der Linden, E.C.M. (2002) I-Type lectins. *Biochim. Biophys. Acta* **1572**, 294–316 https://doi.org/10.1016/S0304-4165(02)00316-1

134 Reutter, W., Stäsche, R., Stehling, P. and Baum, O. (1997) The biology of sialic acids: insights into their structure, metabolism and function in particular during viral infection. In *Glycosciences: Status and Perspectives* (Gabius, H.-J. and Gabius, S., eds), pp. 245–259, Chapman & Hall, London, Weinheim

135 Joziasse, D.H., Schiphorst, W.E.C.M., van den Eijnden, D.H., van Kuik, J.A., van Halbeek, H. and Vliegenthart, J.F.G. (1987) Branch specificity of bovine colostrum CMP-sialic acid: Galβ1-4GlcNAc-R α2-6-sialyltransferase. Sialylation of bi-, tri-, and tetraantennary oligosaccharides and glycopeptides of the *N*-acetyllactosamine type. *J. Biol. Chem.* **262**, 2025–2033 PMID:3546284

136 Dahms, N.M. and Hancock, M.K. (2002) P-Type lectins. *Biochim. Biophys. Acta* **1572**, 317–340 https://doi.org/10.1016/S0304-4165(02)00317-3

137 Kornfeld, S. (2018) A lifetime of adventures in glycobiology. *Annu. Rev. Biochem.* **87**, 1–21 https://doi.org/10.1146/annurev-biochem-062917-011911

138 Iida, S.-i., Yamamoto, K. and Irimura, T. (1999) Interaction of human macrophage C-type lectin with *O*-linked *N*-acetylgalactosamine residues on mucin glycopeptides. *J. Biol. Chem.* **274**, 10697–10705 https://doi.org/10.1074/jbc.274.16.10697

139 Ju, T., Otto, V.I. and Cummings, R.D. (2011) The Tn antigen: structural simplicity and biological complexity. *Angew. Chem. Int. Ed. Engl.* **50**, 1770–1791 https://doi.org/10.1002/anie.201002313

140 Mortezai, N., Behnken, H.N., Kurze, A.K., Ludewig, P., Buck, F., Meyer, B. et al. (2013) Tumor-associated Neu5Ac-Tn and Neu5Gc-Tn antigens bind to C-type lectin CLEC10A (CD301, MGL). *Glycobiology* **23**, 844–852 https://doi.org/10.1093/glycob/cwt021

141 Brinkman-van der Linden, E.C.M. and Varki, A. (2000) New aspects of siglec binding specificities, including the significance of fucosylation and of the sialyl-T_n epitope. *J. Biol. Chem.* **275**, 8625–8632 https://doi.org/10.1074/jbc.275.12.8625

142 Angata, T., Tabuchi, Y., Nakamura, K. and Nakamura, M. (2007) Siglec-15: an immune system Siglec conserved throughout vertebrate evolution. *Glycobiology* **17**, 838–846 https://doi.org/10.1093/glycob/cwm049

143 Takamiya, R., Ohtsubo, K., Takamatsu, S., Taniguchi, N. and Angata, T. (2013) The interaction between Siglec-15 and tumor-associated sialyl-Tn antigen enhances TGF-β secretion from monocytes/macrophages through the DAP12-Syk pathway. *Glycobiology* **23**, 178–187 https://doi.org/10.1093/glycob/cws139

144 Wang, J., Shiratori, I., Satoh, T., Lanier, L.L. and Arase, H. (2008) An essential role of sialylated *O*-linked sugar chains in the recognition of mouse CD99 by paired Ig-like type 2 receptor (PILR). *J. Immunol.* **180**, 1686–1693 https://doi.org/10.4049/jimmunol.180.3.1686

145 Lu, Q., Lu, G., Qi, J., Wang, H., Xuan, Y., Wang, Q. et al. (2014) PILRα and PILRβ have a siglec fold and provide the basis of binding to sialic acid. *Proc. Natl Acad. Sci. U.S.A.* **111**, 8221–8226 https://doi.org/10.1073/pnas.1320716111

146 Furukawa, A., Kakita, K., Yamada, T., Ishizuka, M., Sakamoto, J., Hatori, N. et al. (2017) Structural and thermodynamic analyses reveal critical features of glycopeptide recognition by the human PILRα immune cell receptor. *J. Biol. Chem.* **292**, 21128–21136 https://doi.org/10.1074/jbc.M117.799239

147 Dam, T.K. and Brewer, C.F. (2010) Multivalent lectin-carbohydrate interactions energetics and mechanisms of binding. *Adv. Carbohydr. Chem. Biochem.* **63**, 139–164 https://doi.org/10.1016/S0065-2318(10)63005-3

148 Krzeminski, M., Singh, T., André, S., Lensch, M., Wu, A.M., Bonvin, A.M.J.J. et al. (2011) Human galectin-3 (Mac-2 antigen): defining molecular switches of affinity to natural glycoproteins, structural and dynamic aspects of glycan binding by flexible ligand docking and putative regulatory sequences in the proximal promoter region. *Biochim. Biophys. Acta* **1810**, 150–161 https://doi.org/10.1016/j.bbagen.2010.11.001

149 Murugaesu, N., Iravani, M., van Weverwijk, A., Ivetic, A., Johnson, D.A., Antonopoulos, A. et al. (2014) An *in vivo* functional screen identifies ST6GalNAc2 sialyltransferase as a breast cancer metastasis suppressor. *Cancer Discov.* **4**, 304–317 https://doi.org/10.1158/2159-8290.CD-13-0287

150 Moonens, K. and Remaut, H. (2017) Evolution and structural dynamics of bacterial glycan binding adhesins. *Curr. Opin. Struct. Biol.* **44**, 48–58 https://doi.org/10.1016/j.sbi.2016.12.003

151 Tzeng, S.F., Tsai, C.H., Chao, T.K., Chou, Y.C., Yang, Y.C., Tsai, M.H. et al. (2018) *O*-Glycosylation-mediated signaling circuit drives metastatic castration-resistant prostate cancer. *FASEB J.* **32**, 6869–6882 https://doi.org/10.1096/fj.201800687

152 Klaas, M. and Crocker, P.R. (2012) Sialoadhesin in recognition of self and non-self. *Sem. Immunopathol.* **34**, 353–364 https://doi.org/10.1007/s00281-012-0310-3

153 Beatson, R., Tajadura-Ortega, V., Achkova, D., Picco, G., Tsourouktsoglou, T.D., Klausing, S. et al. (2016) The mucin MUC1 modulates the tumor immunological microenvironment through engagement of the lectin Siglec-9. *Nat. Immunol.* **17**, 1273–1281 https://doi.org/10.1038/ni.3552

154 Blixt, O., Collins, B.E., van den Nieuwenhof, I.M., Crocker, P.R. and Paulson, J.C. (2003) Sialoside specificity of the siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin-associated glycoprotein. *J. Biol. Chem.* **278**, 31007–31019 https://doi.org/10.1074/jbc.M304331200

155 Vyas, A.A., Blixt, O., Paulson, J.C. and Schnaar, R.L. (2005) Potent glycan inhibitors of myelin-associated glycoprotein enhance axon outgrowth *in vitro*. *J. Biol. Chem.* **280**, 16305–16310 https://doi.org/10.1074/jbc.M500250200

156 Tsuboi, S., Sutoh, M., Hatakeyama, S., Hiraoka, N., Habuchi, T., Horikawa, Y. et al. (2011) A novel strategy for evasion of NK cell immunity by tumours expressing core 2 *O*-glycans. *EMBO J.* **30**, 3173–3185 https://doi.org/10.1038/emboj.2011.215

157 Valenzuela, H.F., Pace, K.E., Cabrera, P.V., White, R., Porvari, K., Kaija, H. et al. (2007) *O*-Glycosylation regulates LNCaP prostate cancer cell susceptibility to apoptosis induced by galectin-1. *Cancer Res.* **67**, 6155–6162 https://doi.org/10.1158/0008-5472.CAN-05-4431

158 Hiraoka, N., Kawashima, H., Petryniak, B., Nakayama, J., Mitoma, J., Marth, J.D. et al. (2004) Core 2 branching β1,6-*N*-acetylglucosaminyltransferase and high endothelial venule-restricted sulfotransferase collaboratively control lymphocyte homing. *J. Biol. Chem.* **279**, 3058–3067 https://doi.org/10.1074/jbc.M311150200
159 Kumar, R., Camphausen, R.T., Sullivan, F.X. and Cumming, D.A. (1996) Core 2 β-1,6-*N*-acetylglucosaminyltransferase enzyme activity is critical for P-selectin glycoprotein ligand-1 binding to P-selectin. *Blood* **88**, 3872–3879 PMID:8916952
160 Smith, M.J., Smith, B.R., Lawrence, M.B. and Snapp, K.R. (2004) Functional analysis of the combined role of the *O*-linked branching enzyme core 2 β1,6-*N*-glucosaminyltransferase and dimerization of P-selectin glycoprotein ligand-1 in rolling on P-selectin. *J. Biol. Chem.* **279**, 21984–21991 https://doi.org/10.1074/jbc.M402731200
161 Somers, W.S., Tang, J., Shaw, G.D. and Camphausen, R.T. (2000) Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to sLe^x and PSGL-1. *Cell* **103**, 467–479 https://doi.org/10.1016/S0092-8674(00)00138-0
162 Becker, D.J. and Lowe, J.B. (2003) Fucose: biosynthesis and biological functions in mammals. *Glycobiology* **13**, 41R–53R https://doi.org/10.1093/glycob/cwg054
163 Lühn, K. and Wild, M.K. (2012) Human deficiencies of fucosylation and sialylation affecting selectin ligands. *Semin. Immunopathol.* **34**, 383–399 https://doi.org/10.1007/s00281-012-0304-1
164 Hennet, T. and Cabalzar, J. (2015) Congenital disorders of glycosylation: a concise chart of glycocalyx dysfunction. *Trends Biochem. Sci.* **40**, 377–384 https://doi.org/10.1016/j.tibs.2015.03.002
165 Nagae, M. and Yamaguchi, Y. (2015) Sugar recognition and protein-protein interaction of mammalian lectins conferring diverse functions. *Curr. Opin. Struct. Biol.* **34**, 108–115 https://doi.org/10.1016/j.sbi.2015.08.005
166 Nakakita, S., Itoh, A., Nakakita, Y., Nonaka, Y., Ogawa, T., Nakamura, T. et al. (2016) Cooperative interactions of oligosaccharide and peptide moieties of a glycopeptide derived from IgE with galectin-9. *J. Biol. Chem.* **291**, 968–979 https://doi.org/10.1074/jbc.M115.694448
167 Varki, A., Hooshmand, F., Diaz, S., Varki, N.M. and Hedrick, S.M. (1991) Developmental abnormalities in transgenic mice expressing a sialic acid-specific 9-*O*-acetyltransesterase. *Cell* **65**, 65–74 https://doi.org/10.1016/0092-8674(91)90408-Q
168 Langereis, M.A., Bakkers, M.J., Deng, L., Padler-Karavani, V., Vervoort, S.J., Hulswit, R.J. et al. (2015) Complexity and diversity of the mammalian sialome revealed by nidovirus virolectins. *Cell Rep.* **11**, 1966–1978 https://doi.org/10.1016/j.celrep.2015.05.044
169 Mandal, C., Schwartz-Albiez, R. and Vlasak, R. (2015) Functions and biosynthesis of *O*-acetylated sialic acids. *Top. Curr. Chem.* **366**, 1–30 https://doi.org/10.1007/128_2011_310
170 Holdener, B.C. and Haltiwanger, R.S. (2019) Protein *O*-fucosylation: structure and function. *Curr. Opin. Struct. Biol.* **56**, 78–86 https://doi.org/10.1016/j.sbi.2018.12.005
171 Luca, V.C., Jude, K.M., Pierce, N.W., Nachury, M.V., Fischer, S. and Garcia, K.C. (2015) Structural basis for Notch1 engagement of Delta-like 4. *Science* **347**, 847–853 https://doi.org/10.1126/science.1261093
172 Luca, V.C., Kim, B.C., Ge, C., Kakuda, S., Wu, D., Roein-Peikar, M. et al. (2017) Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. *Science* **355**, 1320–1324 https://doi.org/10.1126/science.aaf9739
173 Dobson, C.M., Hempel, S.J., Stalnaker, S.H., Stuart, R. and Wells, L. (2013) *O*-Mannosylation and human disease. *Cell. Mol. Life Sci.* **70**, 2849–2857 https://doi.org/10.1007/s00018-012-1193-0
174 Manya, H. (2011) Biosynthetic pathway of *O*-mannosyl glycan in mammals. *Trends Glycosci. Glycotechnol.* **23**, 272–283 https://doi.org/10.4052/tigg.23.272
175 Praissman, J.L. and Wells, L. (2014) Mammalian *O*-mannosylation pathway: glycan structures, enzymes, and protein substrates. *Biochemistry* **53**, 3066–3078 https://doi.org/10.1021/bi500153y
176 Hohenester, E. (2019) Laminin G-like domains: dystroglycan-specific lectins. *Curr. Opin. Struct. Biol.* **56**, 56–63 https://doi.org/10.1016/j.sbi.2018.11.007
177 Buddecke, E. (2009) Proteoglycans. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 199–216, Wiley-VCH, Weinheim, Germany
178 Gahmberg, C.G. and Tolvanen, M. (1996) Why mammalian cell surface proteins are glycoproteins. *Trends Biochem. Sci.* **21**, 308–311 https://doi.org/10.1016/S0968-0004(96)10034-7
179 Thudichum, J.L.W. (1874) Researches on the chemical constitution of the brain. In *Report of the Medical Offices of the Privy Council and Local Government Board*, pp. 113–247, Eyre and Spottiswoode for H.M.S.O., London, U.K.
180 Thudichum, J.L.W. (1884) *A Treatise on the Chemical Constitution of the Brain*. Baillière, Tindall & Cox, London
181 Klenk, E. (1942) Über die Ganglioside, eine neue Gruppe von zuckerhaltigen Gehirnlipoiden. *Hoppe Seyler's Z. Physiol. Chem.* **273**, 76–86 https://doi.org/10.1515/bchm2.1942.273.1-2.76
182 Klenk, E. (1970) On the discovery and chemistry of neuraminic acid and gangliosides. *Chem. Phys. Lipids* **5**, 193–197 https://doi.org/10.1016/0009-3084(70)90017-4
183 Merrill, Jr, A.H., Wang, M.D., Park, M. and Sullards, M.C. (2007) (Glyco)sphingolipidology: an amazing challenge and opportunity for systems biology. *Trends Biochem. Sci.* **32**, 457–468 https://doi.org/10.1016/j.tibs.2007.09.004
184 Yu, R.K., Yanagisawa, M. and Ariga, T. (2007) Glycosphingolipid structures. In *Comprehensive Glycoscience* (Kamerling, J.P., ed.), Elsevier, Oxford, U.K.
185 Kopitz, J. (2009) Glycolipids. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 177–198, Wiley-VCH, Weinheim, Germany
186 Kopitz, J. (2017) Lipid glycosylation: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* **147**, 175–198 https://doi.org/10.1007/s00418-016-1518-4
187 Ledeen, R.W. and Wu, G. (2009) Neurobiology meets glycosciences. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 495–516, Wiley-VCH, Weinheim, Germany
188 Schengrund, C.-L. (2015) Gangliosides: glycosphingolipids essential for normal neural development and function. *Trends Biochem. Sci.* **40**, 397–406 https://doi.org/10.1016/j.tibs.2015.03.007
189 Gill, D.M. (1976) The arrangement of subunits in cholera toxin. *Biochemistry* **15**, 1242–1248 https://doi.org/10.1021/bi00651a011
190 van Heyningen, S. (1982) Cholera toxin. *Biosci. Rep.* **2**, 135–146 https://doi.org/10.1007/BF01116376
191 Honke, K., Tsuda, M., Hirahara, Y., Ishii, A., Makita, A. and Wada, Y. (1997) Molecular cloning and expression of cDNA encoding human 3′-phosphoadenylylsulfate:galactosylceramide 3′-sulfotransferase. *J. Biol. Chem.* **272**, 4864–4868 https://doi.org/10.1074/jbc.272.8.4864

Biochemical Journal (2019) **476** 2623–2655
https://doi.org/10.1042/BCJ20170853

192 Fukuda, M., Hiraoka, N., Akama, T.O. and Fukuda, M.N. (2001) Carbohydrate-modifying sulfotransferases: structure, function, and pathophysiology. *J. Biol. Chem.* **276**, 47747–47750 https://doi.org/10.1074/jbc.R100049200
193 Hemmerich, S., Verdugo, D. and Rath, V.L. (2004) Strategies for drug discovery by targeting sulfation pathways. *Drug Discov. Today* **9**, 967–975 https://doi.org/10.1016/S1359-6446(04)03261-1
194 Roberts, D.D., Rao, C.N., Magnani, J.L., Spitalnik, S.L., Liotta, L.A. and Ginsburg, V. (1985) Laminin binds specifically to sulfated glycolipids. *Proc. Natl Acad. Sci. U.S.A.* **82**, 1306–1310 https://doi.org/10.1073/pnas.82.5.1306
195 Kobayashi, T., Honke, K., Kuramitsu, Y., Hosokawa, M., Miyazaki, T., Murata, J. et al. (1994) Cell-surface sulfoglycolipids are involved in the attachment of renal-cancer cells to laminin. *Int. J. Cancer* **56**, 281–285 https://doi.org/10.1002/ijc.2910560223
196 Aruffo, A., Kolanus, W., Walz, G., Fredman, P. and Seed, B. (1991) CD62/P-selectin recognition of myeloid and tumor cell sulfatides. *Cell* **67**, 35–44 https://doi.org/10.1016/0092-8674(91)90570-O
197 Suzuki, Y., Toda, Y., Tamatani, T., Watanabe, T., Suzuki, T., Nakao, T. et al. (1993) Sulfated glycolipids are ligands for a lymphocyte homing receptor, L-selectin (LECAM-1), binding epitope in sulfated sugar chain. *Biochem. Biophys. Res. Commun.* **190**, 426–434 https://doi.org/10.1006/bbrc.1993.1065
198 García, J., Callewaert, N. and Borsig, L. (2007) P-Selectin mediates metastatic progression through binding to sulfatides on tumor cells. *Glycobiology* **17**, 185–196 https://doi.org/10.1093/glycob/cwl059
199 Miura, R., Aspberg, A., Ethell, I.M., Hagihara, K., Schnaar, R.L., Ruoslahti, E. et al. (1999) The proteoglycan lectin domain binds sulfated cell surface glycolipids and promotes cell adhesion. *J. Biol. Chem.* **274**, 11431–11438 https://doi.org/10.1074/jbc.274.16.11431
200 Phongsisay, V., Iizasa, E., Hara, H. and Yamasaki, S. (2015) 3-*O*-Sulfo-β-D-galactose moiety of endogenous sulfoglycolipids is a potential ligand for immunoglobulin-like receptor LMIR5. *Mol. Immunol.* **63**, 595–599 https://doi.org/10.1016/j.molimm.2014.07.023
201 Danielsen, E.M. and van Deurs, B. (1997) Galectin-4 and small intestinal brush border enzymes form clusters. *Mol. Biol. Cell* **8**, 2241–2251 https://doi.org/10.1091/mbc.8.11.2241
202 Delacour, D., Gouyer, V., Zanetta, J.-P., Drobecq, H., Leteurtre, E., Grard, G. et al. (2005) Galectin-4 and sulfatides in apical membrane trafficking in enterocyte-like cells. *J. Cell Biol.* **169**, 491–501 https://doi.org/10.1083/jcb.200407073
203 Morelle, W., Stechly, L., André, S., van Seuningen, I., Porchet, N., Gabius, H.-J. et al. (2009) Glycosylation pattern of brush border-associated glycoproteins in enterocyte-like cells: involvement of complex-type *N*-glycans in apical trafficking. *Biol. Chem.* **390**, 529–544 https://doi.org/10.1515/BC.2009.075
204 Stechly, L., Morelle, W., Dessein, A.F., André, S., Grard, G., Trinel, D. et al. (2009) Galectin-4-regulated delivery of glycoproteins to the brush border membrane of enterocyte-like cells. *Traffic* **10**, 438–450 https://doi.org/10.1111/j.1600-0854.2009.00882.x
205 Velasco, S., Díez-Revuelta, N., Hernández-Iglesias, T., Kaltner, H., André, S., Gabius, H.-J. et al. (2013) Neuronal galectin-4 is required for axon growth and for the organization of axonal membrane L1 delivery and clustering. *J. Neurochem.* **125**, 49–62 https://doi.org/10.1111/jnc.12148
206 Attrill, H., Imamura, A., Sharma, R.S., Kiso, M., Crocker, P.R. and van Aalten, D.M. (2006) Siglec-7 undergoes a major conformational change when complexed with the α(2,8)-disialylganglioside GT1b. *J. Biol. Chem.* **281**, 32774–32783 https://doi.org/10.1074/jbc.M601714200
207 Kawasaki, Y., Ito, A., Withers, D.A., Taima, T., Kakoi, N., Saito, S. et al. (2010) Ganglioside DSGb5, preferred ligand for Siglec-7, inhibits NK cell cytotoxicity against renal cell carcinoma cells. *Glycobiology* **20**, 1373–1379 https://doi.org/10.1093/glycob/cwq116
208 Malaker, S.A., Pedram, K., Ferracane, M.J., Bensing, B.A., Krishnan, V., Pett, C. et al. (2019) The mucin-selective protease stcE enables molecular and functional analysis of human cancer-associated mucins. *Proc. Natl Acad. Sci. U.S.A.* **116**, 7278–7287 https://doi.org/10.1073/pnas.1813020116
209 Vyas, A.A., Patel, H.V., Fromholt, S.E., Heffer-Lauc, M., Vyas, K.A., Dang, J. et al. (2002) Gangliosides are functional nerve cell ligands for myelin-associated glycoprotein (MAG), an inhibitor of nerve regeneration. *Proc. Natl Acad. Sci. U.S.A.* **99**, 8412–8417 https://doi.org/10.1073/pnas.072211699
210 Schnaar, R.L. (2010) Brain gangliosides in axon-myelin stability and axon regeneration. *FEBS Lett.* **584**, 1741–1747 https://doi.org/10.1016/j.febslet.2009.10.011
211 Pan, B., Fromholt, S.E., Hess, E.J., Crawford, T.O., Griffin, J.W., Sheikh, K.A. et al. (2005) Myelin-associated glycoprotein and complementary axonal ligands, gangliosides, mediate axon stability in the CNS and PNS: neuropathology and behavioral deficits in single- and double-null mice. *Exp. Neurol.* **195**, 208–217 https://doi.org/10.1016/j.expneurol.2005.04.017
212 Mehta, N.R., Nguyen, T., Bullen, Jr, J.W., Griffin, J.W. and Schnaar, R.L. (2010) Myelin-associated glycoprotein (MAG) protects neurons from acute toxicity using a ganglioside-dependent mechanism. *ACS Chem. Neurosci.* **1**, 215–222 https://doi.org/10.1021/cn900029p
213 Pronker, M.F., Lemstra, S., Snijder, J., Heck, A.J.R., Thies-Weesie, D.M., Pasterkamp, R.J. et al. (2016) Structural basis of myelin-associated glycoprotein adhesion and signalling. *Nat. Commun.* **7**, 13584 https://doi.org/10.1038/ncomms13584
214 Schengrund, C.L. and Repman, M.A. (1982) Density-dependent changes in gangliosides and sialidase activity of murine neuroblastoma cells. *J. Neurochem.* **39**, 940–947 https://doi.org/10.1111/j.1471-4159.1982.tb11480.x
215 Kopitz, J., von Reitzenstein, C., Mühl, C. and Cantz, M. (1994) Role of plasma membrane ganglioside sialidase of human neuroblastoma cells in growth control and differentiation. *Biochem. Biophys. Res. Commun.* **199**, 1188–1193 https://doi.org/10.1006/bbrc.1994.1356
216 Kopitz, J., Mühl, C., Ehemann, V., Lehmann, C. and Cantz, M. (1997) Effects of cell surface ganglioside sialidase inhibition on growth control and differentiation of human neuroblastoma cells. *Eur. J. Cell Biol.* **73**, 1–7 PMID:9174666
217 Wu, G. and Ledeen, R.W. (1991) Stimulation of neurite outgrowth in neuroblastoma cells by neuraminidase: putative role of GM1 ganglioside in differentiation. *J. Neurochem.* **56**, 95–104 https://doi.org/10.1111/j.1471-4159.1991.tb02567.x
218 Siebert, H.-C., André, S., Lu, S.-Y., Frank, M., Kaltner, H., van Kuik, J.A. et al. (2003) Unique conformer selection of human growth-regulatory lectin galectin-1 for ganglioside GM1 versus bacterial toxins. *Biochemistry* **42**, 14762–14773 https://doi.org/10.1021/bi035477c
219 Kopitz, J., von Reitzenstein, C., Burchert, M., Cantz, M. and Gabius, H.-J. (1998) Galectin-1 is a major receptor for ganglioside GM1, a product of the growth-controlling activity of a cell surface ganglioside sialidase, on human neuroblastoma cells in culture. *J. Biol. Chem.* **273**, 11205–11211 https://doi.org/10.1074/jbc.273.18.11205
220 Kopitz, J., von Reitzenstein, C., André, S., Kaltner, H., Uhl, J., Ehemann, V. et al. (2001) Negative regulation of neuroblastoma cell growth by carbohydrate-dependent surface binding of galectin-1 and functional divergence from galectin-3. *J. Biol. Chem.* **276**, 35917–35923 https://doi.org/10.1074/jbc.M105135200

221 Wu, G., Lu, Z.H., André, S., Gabius, H.-J. and Ledeen, R.W. (2016) Functional interplay between ganglioside GM1 and cross-linking galectin-1 induces axon-like neuritogenesis via integrin-based signaling and TRPC5-dependent Ca^{2+} influx. *J. Neurochem.* **136**, 550–563 https://doi.org/10.1111/jnc.13418
222 Wang, J., Lu, Z.H., Gabius, H.-J., Rohowsky-Kochan, C., Ledeen, R.W. and Wu, G. (2009) Cross-linking of GM1 ganglioside by galectin-1 mediates regulatory T cell activity involving TRPC5 channel activation: possible role in suppressing experimental autoimmune encephalomyelitis. *J. Immunol.* **182**, 4036–4045 https://doi.org/10.4049/jimmunol.0802981
223 Ledeen, R.W., Wu, G., André, S., Bleich, D., Huet, G., Kaltner, H. et al. (2012) Beyond glycoproteins as galectin counterreceptors: tumor/effector T cell growth control via ganglioside GM1. *Ann. N. Y. Acad. Sci.* **1253**, 206–221 https://doi.org/10.1111/j.1749-6632.2012.06479.x
224 Ledeen, R.W., Kopitz, J., Abad-Rodriguez, J. and Gabius, H.J. (2018) Glycan chains of gangliosides: functional ligands for tissue lectins (siglecs/galectins). *Progr. Mol. Biol. Transl. Sci.* **156**, 289–324 https://doi.org/10.1016/bs.pmbts.2017.12.004
225 Kalka, D., von Reitzenstein, C., Kopitz, J. and Cantz, M. (2001) The plasma membrane ganglioside sialidase cofractionates with markers of lipid rafts. *Biochem. Biophys. Res. Commun.* **283**, 989–993 https://doi.org/10.1006/bbrc.2001.4864
226 Ledeen, R.W. and Wu, G. (2015) The multi-tasked life of GM1 ganglioside, a true factotum of nature. *Trends Biochem. Sci.* **40**, 407–418 https://doi.org/10.1016/j.tibs.2015.04.005
227 Ichikawa, N., Iwabuchi, K., Kurihara, H., Ishii, K., Kobayashi, T., Sasaki, T. et al. (2009) Binding of laminin-1 to monosialoganglioside GM1 in lipid rafts is crucial for neurite outgrowth. *J. Cell Sci.* **122**, 289–299 https://doi.org/10.1242/jcs.030338
228 Yoon, S.J., Nakayama, K., Hikita, T., Handa, K. and Hakomori, S.I. (2006) Epidermal growth factor receptor tyrosine kinase is modulated by GM3 interaction with N-linked GlcNAc termini of the receptor. *Proc. Natl Acad. Sci. U.S.A.* **103**, 18987–18991 https://doi.org/10.1073/pnas.0609281103
229 Bucior, I., Burger, M.M. and Fernàndez-Busquets, X. (2009) Carbohydrate-carbohydrate interactions. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 347–362, Wiley-VCH, Weinheim, Germany
230 Lee, Y.C. (1992) Biochemistry of carbohydrate-protein interaction. *FASEB J.* **6**, 3193–3200 https://doi.org/10.1096/fasebj.6.13.1397841
231 Lee, R.T. and Lee, Y.C. (1994) Enhanced biochemical affinities of multivalent neoglycoconjugates. In *Neoglycoconjugates. Preparation and Applications* (Lee, Y.C. and Lee, R.T., eds), pp. 23–50, Academic Press, San Diego
232 Fraser, I.P., Koziel, H. and Ezekowitz, R.A.B. (1998) The serum mannose-binding protein and the macrophage mannose receptor are pattern recognition molecules that link innate and adaptive immunity. *Sem. Immunol.* **10**, 363–372 https://doi.org/10.1006/smim.1998.0141
233 Gupta, G. and Surolia, A. (2007) Collectins: sentinels of innate immunity. *BioEssays* **29**, 452–464 https://doi.org/10.1002/bies.20573
234 Yamaguchi, Y. (2000) Lecticans: organizers of the brain extracellular matrix. *Cell. Mol. Life Sci.* **57**, 276–289 https://doi.org/10.1007/PL00000690
235 Iozzo, R.V. and Murdoch, A.D. (1996) Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J.* **10**, 598–614 https://doi.org/10.1096/fasebj.10.5.8621059
236 Viapiano, M.S. and Matthews, R.T. (2006) From barriers to bridges: chondroitin sulfate proteoglycans in neuropathology. *Trends Mol. Med.* **12**, 488–496 https://doi.org/10.1016/j.molmed.2006.08.007
237 Thomas, W.E., Vogel, V. and Sokurenko, E. (2008) Biophysics of catch bonds. *Annu. Rev. Biophys.* **37**, 399–416 https://doi.org/10.1146/annurev.biophys.37.032807.125804
238 Dulal, H.P., Adachi, Y., Ohno, N. and Yamaguchi, Y. (2018) β-Glucan-induced cooperative oligomerization of Dectin-1 C-type lectin-like domain. *Glycobiology* **28**, 612–623 https://doi.org/10.1093/glycob/cwy039
239 Flores-Ibarra, A., Vértesy, S., Medrano, F.J., Gabius, H.-J. and Romero, A. (2018) Crystallization of a human galectin-3 variant with two ordered segments in the shortened N-terminal tail. *Sci. Rep.* **8**, 9835 https://doi.org/10.1038/s41598-018-28235-x
240 Weinmann, D., Kenn, M., Schmidt, S., Schmidt, K., Walzer, S.M., Kubista, B. et al. (2018) Galectin-8 induces functional disease markers in human osteoarthritis and cooperates with galectins-1 and -3. *Cell. Mol. Life Sci.* **75**, 4187–4205 https://doi.org/10.1007/s00018-018-2856-2
241 Manning, J.C., García Caballero, G., Knospe, C., Kaltner, H. and Gabius, H.-J. (2017) Network analysis of adhesion/growth-regulatory galectins and their binding sites in adult chicken retina and choroid. *J. Anat.* **231**, 23–37 https://doi.org/10.1111/joa.12612
242 Manning, J.C., García Caballero, G., Knospe, C., Kaltner, H. and Gabius, H.-J. (2018) Three-step monitoring of glycan and galectin profiles in the anterior segment of the adult chicken eye. *Ann. Anat.* **217**, 66–81 https://doi.org/10.1016/j.aanat.2018.02.002
243 Sanchez-Ruderisch, H., Fischer, C., Detjen, K.M., Welzel, M., Wimmel, A., Manning, J.C. et al. (2010) Tumor suppressor $p16^{INK4a}$: downregulation of galectin-3, an endogenous competitor of the pro-anoikis effector galectin-1, in a pancreatic carcinoma model. *FEBS J.* **277**, 3552–3563 https://doi.org/10.1111/j.1742-4658.2010.07764.x
244 Ludwig, A.-K., Kaltner, H., Kopitz, J. and Gabius, H.-J. (2019) Lectinology 4.0: altering modular (ga)lectin display for functional analysis and biomedical applications. *Biochim. Biophys. Acta* **1863**, 935–940 https://doi.org/10.1016/j.bbagen.2019.03.005
245 Tasumi, S. and Vasta, G.R. (2007) A galectin of unique domain organization from hemocytes of the Eastern oyster (*Crassostrea virginica*) is a receptor for the protistan parasite *Perkinsus marinus*. *J. Immunol.* **179**, 3086–3098 https://doi.org/10.4049/jimmunol.179.5.3086
246 Feng, C., Ghosh, A., Amin, M.N., Giomarelli, B., Shridhar, S., Banerjee, A. et al. (2013) The galectin CvGal1 from the eastern oyster (*Crassostrea virginica*) binds to blood group A oligosaccharides on the hemocyte surface. *J. Biol. Chem.* **288**, 24394–24409 https://doi.org/10.1074/jbc.M113.476531
247 Kopitz, J., Xiao, Q., Ludwig, A.-K., Romero, A., Michalak, M., Sherman, S.E. et al. (2017) Reaction of a programmable glycan presentation of glycodendrimersomes and cells with engineered human lectins to show the sugar functionality of the cell surface. *Angew. Chem. Int. Ed. Engl.* **56**, 14677–14681 https://doi.org/10.1002/anie.201708237
248 Rüdiger, H. and Gabius, H.-J. (2009) The biochemical basis and coding capacity of the sugar code. In *The Sugar Code. Fundamentals of Glycosciences* (Gabius, H.-J., ed.), pp. 3–13, Wiley-VCH, Weinheim, Germany
249 Hricovini, M., Guerrini, M., Bisio, A., Torri, G., Petitou, M. and Casu, B. (2001) Conformation of heparin pentasaccharide bound to antithrombin III. *Biochem. J.* **359**, 265–272 https://doi.org/10.1042/bj3590265
250 Guerrini, M., Guglieri, S., Beccati, D., Torri, G., Viskov, C. and Mourier, P. (2006) Conformational transitions induced in heparin octasaccharides by binding with antithrombin III. *Biochem. J.* **399**, 191–198 https://doi.org/10.1042/BJ20060656