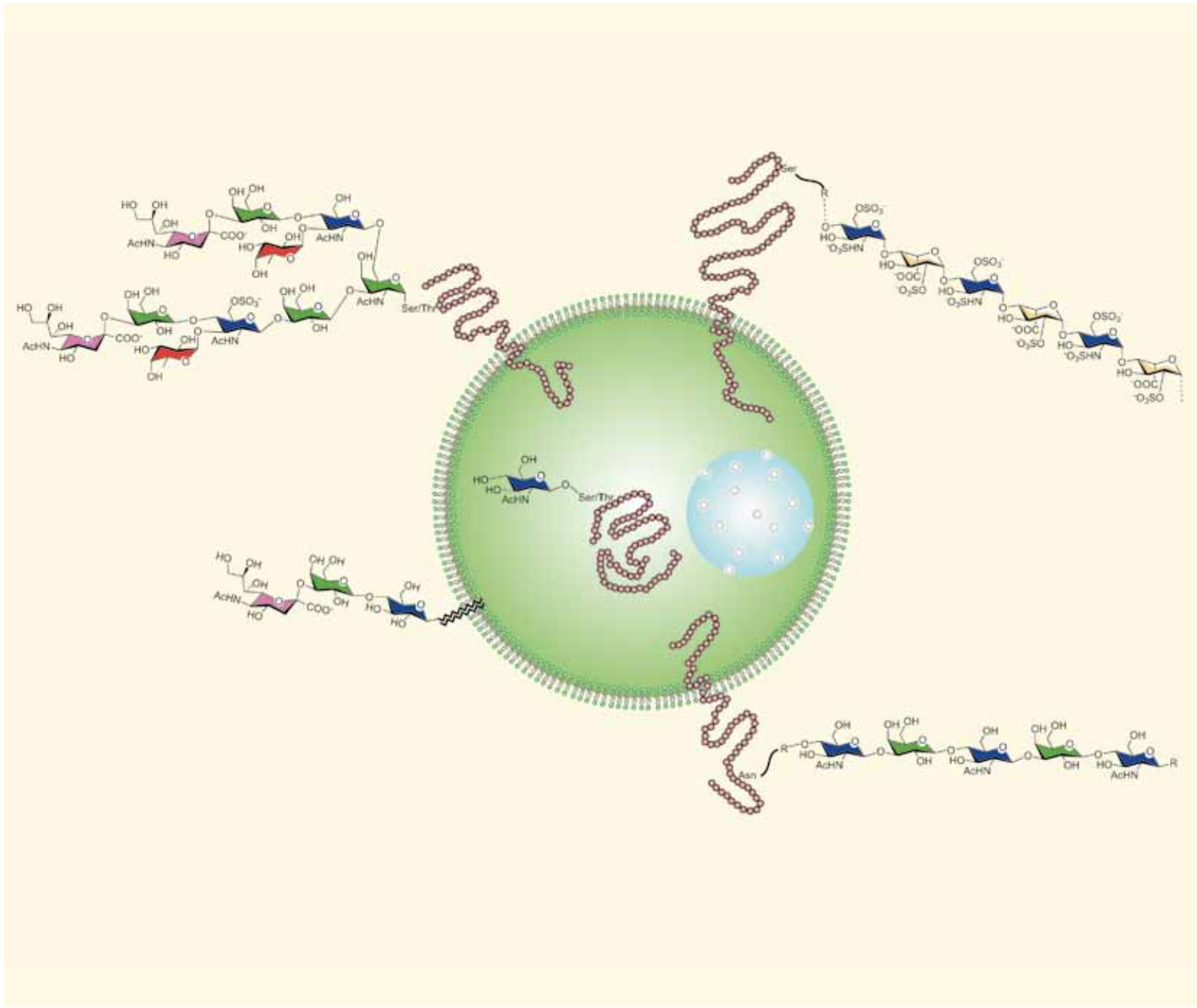
Molecular BioSystems

www.molecularbiosystems.org

Volume 5 | Number 10 | October 2009 | Pages 1077–1248



ISSN 1742 -206X

RSC Publishing

REVIEW
Richard D. Cummings
The repertoire of glycan determinants in the human glycome

PAPER
Phillip C. Wright *et al.*
Systems biology meets synthetic biology: a case study of the metabolic effects of synthetic rewiring

1742-206X(2009)5:10;1-N

The repertoire of glycan determinants in the human glycome

Richard D. Cummings*

Received 21st April 2009, Accepted 11th June 2009
First published as an Advance Article on the web 28th July 2009
DOI: 10.1039/b907931a

The number of glycan determinants that comprise the human glycome is not known. This uncertainty arises from limited knowledge of the total number of distinct glycans and glycan structures in the human glycome, as well as limited information about the glycan determinants recognized by glycan-binding proteins (GBPs), which include lectins, receptors, toxins, microbial adhesins, antibodies, and enzymes. Available evidence indicates that GBP binding sites may accommodate glycan determinants made up of 2 to 6 linear monosaccharides, together with their potential side chains containing other sugars and modifications, such as sulfation, phosphorylation, and acetylation. Glycosaminoglycans, including heparin and heparan sulfate, comprise repeating disaccharide motifs, where a linear sequence of 5 to 6 monosaccharides may be required for recognition. Based on our current knowledge of the composition of the glycome and the size of GBP binding sites, glycoproteins and glycolipids may contain ~3000 glycan determinants with an additional ~4000 theoretical pentasaccharide sequences in glycosaminoglycans. These numbers provide an achievable target for new chemical and/or enzymatic syntheses, and raise new challenges for defining the total glycome and the determinants recognized by GBPs.

Introduction

The diversity of complex carbohydrates in terms of their asymmetry and in their recognition by proteins has been recognized since the initial studies on these macromolecules began over a century ago. As noted in 1894 by Prof. Emil Fischer,[1] "Their [an enzyme's] specific effect on the glucosides might thus be explained by assuming that the intimate contact between the molecules necessary for the release of the chemical reaction is possible only with similar geometrical configurations. To give an illustration I will say that enzyme and glucoside must fit together like lock and key [underline added] in order to be able to exercise a chemical action on each other.... It is an extension of the theory of asymmetry without being a direct consequence of it: for the conviction that the geometrical structure of the molecule even for optical isomers exercises such a great influence on the chemical affinities, in my opinion could only be gained by new actual observations." Prof. Paul Ehrlich, considered the father of immunology, referred to Fischer's perception of the "lock and key" in his own prediction that the specific shapes of molecules are recognized by specific cell receptors,[2] which he restated in his Nobel Prize

Department of Biochemistry, Emory University School of Medicine, 1510 Clifton Rd. #4001, Atlanta, GA 30322, USA. E-mail: rdcummi@emory.edu; Fax: +1-404-727-2738; Tel: +1-404-727-5962

Dr Richard D. Cummings is William Patterson Timmie Professor and Chair of the Department of Biochemistry at Emory University School of Medicine. Prior to joining Emory Dr Cummings was the George Lynn Cross Distinguished Research Professor of Biochemistry and Molecular Biology at the University of Oklahoma Health Sciences Center in Oklahoma City, and also held the Ed Miller Endowed Chair in Molecular Biology. There, he was the founder and Director of the Oklahoma Center for Medical Glycobiology. Dr Cummings received his bachelor's degree in 1974 from the University of Montevallo in Montevallo. Alabama and his Ph.D. in 1980 from The Johns Hopkins University in Baltimore, Maryland under his advisor Dr Stephen Roth. Following that he conducted his post-doctoral fellowship with Dr Stuart Komfeld in Hematology/Oncology at the Washington University School of Medicine in St. Louis, Missouri. Dr Cummings has authored and coauthored over 200 peer-reviewed publications in the field of glycobiology abd is an Editor of the first textbook on glycobiology called "Essentials of Glycobiology"(1st edition, 1999) and the 2nd edition of the textbook (2009). Dr Cummings is a nationally recognized expert in the emerging research field of "functional glycomics" and uses a multi-disciplinary approach to address fundamental mechanisms of vascular inflammation, leukocyte trafficking and turnover, vascular parasitism, and developmental embryology, with special emphasis on the roles of complex carbohydrates (glycoconjugates) and glycan-binding proteins. He has had over 20 graduate students, and over 30 post-doctoral fellows. His group at Emory is also the headquarters for the NIH-funded Consortium for Functional Glycomics (Core H) and the Glycomics Center at Emory University.

address in 1908.[3] Fischer's concept of a lock and key proved to be brilliant, as were most of his other scientific interpretations, and led to his recognition for the first Nobel Prize in Chemistry in 1902. We now know that the chemical structures of complex carbohydrates are recognized by a vast array of proteins, including enzymes and glycan-binding proteins (GBPs).

The region of a complex carbohydrate molecule that is recognized by GBPs may be termed the *glycan determinant*, which is defined as a glycan structure with its associated substitutions and aglycone that is required for the *specific recognition* by a GBP. Partial determinants are components of a glycan determinant that are typically necessary but not sufficient for high affinity interactions of a glycoconjugate with a GBP. A common empirical proof of specific recognition is to show that chemical and structural isomers of a glycan determinant are not as well recognized by the GBP. The glycan–GBP interactions range from relatively low to high affinity, an arbitrary qualification that should be interpreted cautiously in the context of the physiological concentrations of the GBP and the potential glycan it recognizes, as well as in regard to whether potential complex glycan determinants have been identified. Here the term GBP is used in a broad sense to include lectins, antibodies, toxins, and adhesins, with the understanding that the GBP does not cause a catalytic change in the glycan sequence, as through action of a glycosidase or glycosyltransferase. However, under certain conditions where catalytic activity is suppressed, both glycosidases and glycosyltransferases could act as GBPs. Other terms that have been used to describe glycan determinants are oligosaccharide epitope and "glycotope", the latter being introduced by Troy,[4] which may be thought of as an immunological epitope that is carbohydrate-based and recognized by an antibody, but could also have a broader definition as any complex glycan determinant recognized by proteins.

The first described glycan determinants

The first glycan determinants to be defined and structurally elucidated were the blood group antigens. Although Prof. Landsteiner discovered the ABO(H) blood group system early in the 20th century,[5] it was not until 1952 that Prof. Watkins and Prof. Morgan began reporting that they were carbohydrates and subsequently elucidated the structures of blood group A, B, and H antigens[6,7] (Table 1). It is interesting to note that Prof. Watkins and Prof. Morgan and others, including Prof. Kabat,[8] utilized both antibodies and plant and animal lectins to help identify the blood group substances. The lectins included the *Anguilla anguilla* (eel) agglutinin (AAA) and the plant agglutinins from *Lotus tetragonolobus* (LTA), and *Ulex europaeus* (UEA), which all bound the α-linked fucose in the H-antigen. This early use of lectins to help define glycan structures showed that lectins, in a fashion similar to antibodies, had specific recognition for glycans, binding to some and not others, and helped establish the conclusion that glycan structures were "bioinformational", and thus, structurally and functionally distinct. Further studies on antibodies and lectins[9–12] continued to support the general interpretation that glycan structures could be distinguished by their interactions with GBPs, which accelerated the quest to explore the glycome. The glycome may be defined as the entire set of glycans in an organism, which may range from mono- to polysaccharides, either free or present in linkage to aglycone moieties, *e.g.* proteins or lipids. Until the 1980's brought modern mass spectrometry and ^{1}H-NMR approaches, the identification and sequencing of glycans was laborious and costly. As the number of defined glycan structures has grown, so has our understanding of glycan diversity and glycan determinants.

In spite of the advances in glycotechnology through the end of the 20th century, there is much to discover about the complex binding sites within GBPs for large glycan determinants and the number of such determinants. Consider that it was recently noted that "Only 12 out of the more than 500 structures of lectins and carbohydrate-binding domains currently present in the Protein Data Bank [13] have a bound ligand corresponding to at least a pentasaccharide", [quote from ref. 14]. The situation has not dramatically improved in the past couple of years. Most GBP/ligand crystallography utilizes small ligands that usually have low affinity and specificity, but are convenient and available for studies. They provide fundamental insights into aspects of glycan–protein interactions that account for the specific recognition of glycan determinants.[15–24] In addition, there are many recent and excellent studies on plant lectins,[14,25–28] but relatively few studies on animal cell GBPs, which is especially disappointing considering the 53 classes of carbohydrate-binding modules (http://www.cazy.org/fam/acc_CBM.html)[29] and more than 20 classes of lectins expressed in animals,[30,31] along with many dozens of antibodies that bind specific glycan determinants (Table 1). Nevertheless, recent studies on GBPs in complex with larger glycans and even glycan processing enzymes, while difficult, have been enormously revealing in terms of identifying the more complex molecular recognition mechanisms employed by GBPs. Examples of such revealing studies include those on DC-SIGN and SIGN-R,[32] langerin,[33] P-selectin,[34] siglec-7,[35] antibody 2G12 to HIV,[36,37] the P-type lectins,[38,39] the variable lymphocyte receptors,[40] galectin-9,[41] α-mannosidase II,[42] the pregnancy-associated malaria protein VAR2CSA complexed with chondroitin sulfate A,[43] the cathepsin K–chondroitin sulfate complex,[44] and eosinophil-granule major basic protein, a C-type lectin, that binds heparin.[45] Clearly, the continuing challenge in future studies is to have usable quantities of defined and complex glycan determinants to fully explore recognition and the basis for glycan recognition and specificity using X-ray crystallography, NMR, and other biophysical approaches.

Glycan diversity

A tremendous diversity of glycans are thought to make up animal glycomes, and many different glycans are recognized by a growing list of GBPs that interpret what has been termed the "glyco-code".[22,31,46–49] Using the nine monosaccharides found in mammalian glycans, Prof. Laine calculated that 10^{12} different hexasaccharides are theoretically possible,[50] but recognized that many theoretical combinations may not be synthesized by the mammalian glycosylation machinery.

Table 1 Some prominent glycan determinants: evidence for extended glycan recognition

Glycan determinant in N- and O-glycans and glycolipids	Trivial name (and reference)[a]
Galβ1-4GlcNAcβ1-R	Type 2 LN (N-acetyllactosamine)[230,231]
Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-R	Type 2 LN_2 (“I” antigen)[232,233]
Galβ1-4GlcNAcβ1-3(Galβ1-4GlcNAcβ1-6)Galβ1-4GlcNAcβ1-R	“I” antigen[234]
Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-R	“I” antigen[137]
Galβ1-3GlcNAcβ1-R	Type 1 LN[235]
Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ1-R	Type 1 LN_2[236]
Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ1-3Galβ1-3GlcNAcβ1-R	Type 1 LN_3[236]
GalNAcα1-Ser/Thr	Tn antigen[237,238]
Galβ1-3GalNAcα1-Ser/Thr	T (also TF) antigen[239–241]
Fucα1-2Galβ1-3GalNAcα1-Ser/Thr	H antigen (type 3)[242]
GalNAcα1-3(Fucα1-2)Galβ1-3GalNAcα1-Ser/Thr	A antigen (type 3))[243]
NeuAcα2-6GalNAcα1-R	Sialyl Tn antigen[244]
NeuAcα2-3Galβ1-3GalNAcα1-Ser/Thr	Sialyl T antigen[245]
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcα1-Ser/Thr	Di-sialyl T antigen[245]
Galα1-3Galβ1-R	α-Gal antigen[246]
GalNAcβ1-4GlcNAc-R	LacdiNAc (LDN)[247]
GalNAcβ1-4(Fucα1-3)GlcNAc-R	Fucosylated LDN (LDNF)[248]
Galβ1-4(Fucα1-3)GlcNAc-R	Lewis x (Le^x) (SSEA-1)[249,250]
Fucα1-2Galβ1-4(Fucα1-3)GlcNAc-R	Lewis y (Le^y)[251]
Galβ1-3(Fucα1-4)GlcNAc-R	Lewis a (Le^a)[252]
Fucα1-2Galβ1-3(Fucα1-4)GlcNAc-R	Lewis b (Le^b)[253]
NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-R	Sialyl Lewis x (SLe^x)[254]
CyclicNeuAcα2-3Galβ1-4(Fucα1-3)(Su-6)GlcNAcβ1-R[b]	Cyclic sialyl 6-sulfo Lewis x[255]
NeuAcα2-3Galβ1-3(Fucα1-4)GlcNAcβ1-R	Sialyl Lewis a (SLe^a) (CA19-9 antigen)[256]
NeuAcα2-3(Su-6)Galβ1-4(Fucα1-3)GlcNAcβ1-R	6′-Sulfo-sialyl Lewis x (6′-sulfo SLe^x)[257]
NeuAcα2-6Galβ1-4(Su-6)GlcNAcβ1-R	2,6-sialyl-Sulfo-LN (6-sialyl-6-sulfo LN)[258]
NeuAcα2-3Galβ1-4(Fucα1-3)(Su-6)GlcNAcβ1-R	6-Sulfo-sialyl Lewis x (6-sulfo SLe^x)[106]
NeuAcα2-3(Su-6)Galβ1-4(Fucα1-3)(Su-6)GlcNAcβ1-R	6,6′-bisSulfo-Lewis x (6,6′-bissulfo Le^x))[106]
Su-3Galβ1-4(Fucα1-3)GlcNAcβ1-R	3′-Sulfo-Lewis x (3′-sulfo Le^x))[106]
Su-3Galβ1-3(Fucα1-4)GlcNAcβ1-R	3′-Sulfo-Lewis a (3′-Sulfo Le^a) [259]
Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-R	Le^x–Le^x[220]
NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-R	$SDLe^x$[260]
NeuAcα2-3Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-R	VIM-2[222]
NeuAcα2-3Galβ1-4GlcNAc-R	3-Sialyl-LN (type 2)[261,262]
NeuAcα2-6Gal(NAc)β1-R	6-Sialyl-GalNAc or 6-Sialyl LN (type 1 or 2)[263]
Fucα1-2Galβ1-4GlcNAcβ1-3Galβ1-R	Blood group H (type 2)[264]
Galα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-3Galβ1-R	Blood group B (type 2)
GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-R	Blood group A (type 2)
GalNAcα1-3(Fucα1-2)Galβ1-4(Fucα1-3)GlcNAcβ1-R	Blood group A (type 2) (A-Le^y)
GalNAcα1-3(Fucα1-2)Galβ1-3GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-R	Blood group A1 (type 3)[265]
Fucα1-2Galβ1-3GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-R	Blood group A2 (A-associated H type 3)[242]
Fucα1-2Galβ1-3GlcNAcβ1-3Galβ1-R	Blood group H (type 1)[266]
Galα1-3(Fucα1-2)Galβ1-3GlcNAcβ1-R	Blood group B (type 1)[266]
GalNAcα1-3(Fucα1-2)Galβ1-3GlcNAcβ1-R	Blood group A (type 1)[266]
Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ1-R	Blood group H (type 2) (Le^y)[267]
Galα1-3(Fucα1-2)Galβ1-3(Fucα1-4)GlcNAcβ1-R	Blood group B (type 1) (B-Le^y)
GalNAcα1-3(Fucα1-2)Galβ1-3(Fucα1-4)GlcNAcβ1-R	Blood group A (type 1) (A-Le^b)[266,268]
Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ1-R	Lewis b (Le^b)[253]
Su-4GalNAcβ1-4GlcNAc-R	4′-sulfated LDN[223]
NeuAcα2-6GalNAcβ1-4GlcNAc-R	Sialylated LDN[269]
NeuAcα2-3(GalNAcβ1-4)Galβ1-4GlcNAcβ1-R	Sd^a/CT antigen[270,271]
NeuAcα2-8(NeuAcα2-8)$_n$NeuAcα2-3Galβ1-4GlcNAcβ1-R	Polysialic acid[272,273]
(NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-6)(NeuAcα2-3Galβ1-3)GalNAcα1-Ser/Thr	SLe^x Core 2 *O*-glycan glycan [121,219]
P-6-Manα1-2Manα1-3(Manα1-3)(P-6-Manα1-6)Manα1-6Manβ1-4GlcNAcβ-R	Diphosphorylated Man_6[39]
NeuAcα2-3Galβ1-4GlcNAcβ1-2Manα1-Ser/Thr	*O*-Linked mannose[274]
NeuAcα2-3Galβ1-4GlcNAcβ1-2(NeuAcα2-3Galβ1-4GlcNAcβ1-6)Manα1-Ser/Thr	2,6-Branched *O*-mannose[275]
Galβ1-4(Fucα1-3)GlcNAcβ1-2Manα1-Ser/Thr	*O*-Mannose Le^x[276]
Su-3GlcAβ1-3Galβ1-4GlcNAcβ1-R	HNK-1 antigen[276]
Su-3GlcAβ1-3Galβ1-4GlcNAcβ1-2Manα1-Ser/Thr	HNK-1 on *O*-mannose[277]
Glycolipids and glycolipid antigens and glycan determinants	
Glcβ1-Cer	Glucosylceramide[278]
Galβ1-Cer	Galactosylceramide[279,280]
Su-3Galβ1-alkyl-2-acyl-s-glycerol	Seminolipid[281]
Su-3Galβ1-Cer	Sulfatide[282]
Galβ1-4Glcβ1-Cer	Lactosylceramide[283]
Su-3Galβ1-4Galβ1-Cer	Ceramide dihexosyl sulfate[284]
Su-3Galβ1-3GalNAcβ1-4Galβ1-4Glcβ1-Cer	Monosulfated gangliotetraosylceramide[285]
GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Lactotriaosylceramide (Lc_3)[286]
Galα1-4Galβ1-4Glcβ1-Cer	P^k antigen (Gb_3, globotriaosylceramide)[287,288]
Galα1-4Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	P1 antigen[289]
GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Globoside (P antigen) (Gb_4) [288]

Table 1 (*continued*)

Glycan determinant in N- and O-glycans and glycolipids	Trivial name (and reference)[a]
GalNAcβ1-3Galα1-3Galβ1-4Glcβ-Cer	Isoglobotetraosylceramide[290–292]
Su-GalNAcβ1-3Galα1-3Galβ1-4Glcβ1-Cer	Mono-sulfated globopentaosylceramide[293]
Su-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ-Cer	Mono-sulfated globotetraosylceramide[294]
Su-3GalNAcβ1-3Galα1-3Galβ1-4Glcβ-Cer	Sulfo-isogloboside[295]
GalNAcβ1-4(GlcNAcβ1-3)Galβ1-4Glcβ1-Cer	LcGg4[296]
Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Gb_5[297]
Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Paragloboside[264,298]
Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	GL-6 fucosylated (globoH)[299]
Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	SSEA-3[300]
NeuAcα2-3Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	GL-7 globoseries ganglioside (SSEA-4)[299]
NeuAcα2-3GalNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Sialosyl paragloboside[301]
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Disialosyl globopentaosylceramide[302]
GalNAcβ1-3Galα1-3Galβ1-4Glcβ1-Cer	Cytolipin R[303,304]
GalNAcα1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Forssman glycolipid[305]
GalNAcβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	*para*-Forssman glycolipid[306]
Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Blood group H (type 4)[307,308]
Galα1-3(Fucα1-2)Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Blood group B (type 4)[307,308]
GalNAcα1-3(Fucα1-2)Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer	Blood group A (type 4)[307,308]
NeuAcα2-3Galβ1-Cer	GM4[309]
NeuAcα2-3Galβ1-4Glcβ1-Cer	GM3[310]
NeuAcα2-8NeuAcα2-3Galβ1-4Glcβ1-Cer	GD3[311–313]
NeuAcα2-8NeuAcα2-8NeuAcα2-3Galβ1-4Glcβ1-Cer	GT3[314]
9-O-NeuAcα2-8NeuAcα2-3Galβ1-4Glcβ1-Cer	9-*O*-Acetyl GD3[315]
GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GD2[316,317]
GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	GM2[318]
GalNAcβ1-4(NeuGcα2-3)Galβ1-4Glcβ1-Cer	*N*-Glycolyl-GM2[299]
GalNAcβ1-4(NeuAcα2-3)Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Sialopentaosylceramide[319]
Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	GM1[311]
Galβ1-3GalNAcβ1-4Galβ1-4Glcβ1-Cer	Asialo-GM1[320]
NeuAcα2-3Galβ1-3GalNAcβ1-4Galβ1-4Glcβ1-Cer	*cis*GM1 (GM1b)[321]
Fucα1-2Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	2-Fucosyl-GM1[322]
Galα1-3(Fucα1-2)Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	B-GM1[322]
Fucα1-2Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	2-Fucosyl-GD1b[322]
Galα1-3(Fucα1-2)Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	B-GD1b[322]
NeuAcα2-8NeuAcα2-3Galβ1-3GalNAcβ1-4Galβ1-4Glcβ1-Cer	GD1 (GD1c)[323]
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-4Galβ1-4Glcβ1-Cer	GD1α[324]
NeuAcα2-3Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	GD1a[325,326]
Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GD1b[327]
GalNAcβ1-4(NeuAcα2-8NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GT2[328]
NeuAcα2-8NeuAcα2-3Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	GT1a[329]
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	GT1aα[330]
NeuAcα2-3Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GT1b[331]
Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GT1c[328]
NeuAcα2-8NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-Cer	GQ1aα[332]
NeuAcα2-8NeuAcα2-3Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GQ1b[225,333]
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-4(NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GQ1bα[330,332]
NeuAcα2-3Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GQ1c[334]
NeuAcα2-8NeuAcα2-3Galβ1-3GalNAcβ1-4(NeuAcα2-8NeuAcα2-8NeuAcα2-3)-Galβ1-4Glcβ1-Cer	GP1c[328]
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-4(NeuAcα2-8NeuAcα2-8NeuAcα2-3)-Galβ1-4Glcβ1-Cer	GP1cα[332]
NeuAcα2-8NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcβ1-4(NeuAcα2-8NeuAcα2-8NeuAcα2-3)Galβ1-4Glcβ1-Cer	GH1cα[332]
Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Mono-fucosyl LN_5[335]
Galβ1-4GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Mono-fucosyl LN_3[336]
(NeuAcα2-3Galβ1-4GlcNAcβ1-6)(NeuAcα2-3Galβ1-4GlcNAcβ1-3)Galβ1-4GlcNAcβ1-R	Disialyl-branched Type 2[337]
Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ1-3(Fucα1-2)Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Extended tetrafucosyl-Le^b[338]
Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Extended trifucosyl-Le^b[339]
Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Dimeric Le^a[340]
NeuAcα2-3Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Sialyl Le^a glycolipid [341]
NeuAcα2-6Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	Monosialylganglioside LSTb[342]
Su-3GlcAβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer	HNK-expressing SGLPG[315]

[a] Common names and symbols for different glycans are used, based on references provided. Most of the selected references here identify antibodies to the glycan determinants shown, but several also refer to specific recognition by other GBPs, such as toxins and lectins. [b] Su = Sulfate.

Glycan diversity arises not only from differences in monosaccharide composition (Gal *versus* GalNAc, *etc.*), anomeric state (α *versus* β), linkage (1–3 *versus* 1–4, *etc.*), and branching of a monosaccharide, but also from substituted components (phosphate, sulfate, *etc.*) and linkage to their aglycones (peptide, lipid, *etc.*).[51] While the total number of glycans in

the human glycome is not known and may only be approached asymptotically over time, an extremely large number of glycans are known. Over 2000 structures have been described in *N*-glycans, and this diversity is even more complex in terms of specific glycoproteins to which they are attached.[47,52] A few years ago it was estimated that, given the known number of glycosyltransferases in the genome and available glycan structural data, there may be only 500 unique glycan structures.[53] Yet by 2007 it was reported that CarbBank, a database of complex carbohydrate structures, contained 23 118 distinct glycan structures,[54] which is consistent with a 1992 estimate that between 600 and 900 unique glycan structures were being deposited in CarbBank per year.[55] The Consortium for Functional Glycomics (www.functionalglycomics.org/), funded by the National Institutes of General Medical Sciences, lists 7500 entries in the glycan structure database, and the Lipid Maps Consortium (www.lipidmaps.org/), also supported by NIGMS, lists nearly 500 different glycolipids in its database. This illustrates the rapid advance in our understanding of the complexity of the human and rodent glycomes.

The mammalian glycome is generated by several hundred genes encoding glycosyltransferases, sulfotransferases, epimerases, kinases, phosphorylating enzymes, glycosidases, transporters, nucleotide sugar synthetic pathways, *etc*.[56–61] The presence of so many genes regulating glycan biosynthesis is not surprising considering that glycosylation is the most common post-translational modification and over one-half of all proteins probably contain one or more attached mono- or oligo(poly)saccharides.[62–65]

Thus, merely in terms of glycoprotein *N*- and *O*-glycans, the diversity of glycoconjugates in humans and rodents is astonishing. Let us consider several examples. One of the largest glycoproteins in animals is the rat sublingual mucin, which has a mass of $\sim 2.2 \times 10^6$ Da and a carbohydrate content of $\sim$81%. It is estimated to have $\sim$1342 *O*-glycans and $\sim$77 *N*-glycans per molecule.[66] Surprisingly, this large glycoprotein contains a relatively small set of *N*- and *O*-glycans, largely characterized by hybrid-type *N*-glycans and core 4-based *O*-glycans, respectively. By contrast, one of the smallest known glycoproteins is CD24, which is a GPI-anchored glycoprotein (murine and human having only 27 and 32 amino acids, respectively) that binds to the cell adhesion molecule L1 in a manner dependent on α2-3-sialic acid in L1.[67] Mouse brain-derived CD24 is >80% carbohydrate and contains a dizzying array of di-, tri-, and tetraantennary complex-type *N*-glycans, as well as small amounts of high mannose- and hybrid-type *N*-glycans. Additionally, CD24 contains extensive poly-*N*-acetyllactosamine sequences that are polyfucosylated (poly-Le^x), along with blood group H antigen, and curiously is lacking in α2-6-sialic acid.[68] It is noteworthy that the glycan structures of rat sublingual mucin and CD24 are not significantly related. Such unrelated glycosylation among glycoproteins is more the norm than the exception. In fact it is now well accepted that protein glycosylation is typically protein-specific, as exemplified by the glycosylation of lysosomal enzymes,[69–71] glycoprotein hormones,[72,73] CD34,[74,75] PSGL-1,[76–78] IgG,[79,80] erythrocyte glycophorin *versus* band 3 glycoprotein,[81–85] serum transferring,[86–88] fibromodulin,[89] glycodelin,[90] *etc*. A complete list of defined, differentially glycosylated glycoproteins would contain hundreds of examples and is outside the scope of this review. However, in considering the glycan determinants expressed in mammalian glycomes, these examples should provoke sobering reflection on the diversity of glycosylation. Such examples also contradict some common misconceptions of glycoscience, such as the notion that protein glycosylation does not vary remarkably from one glycoprotein to the other and that cell-type specific glycosylation is limited.

These examples draw attention to the enormous variety of glycan structures, the vast challenge of modern glycomics, and the near impossibility of using available approaches and resources to both structurally define all glycans in glycoproteins and identify their linkage sites. Thus, the goal of glycomics to "sequence" the human glycome, *i.e*. defining all glycan structures and their aglycone moieties, may be unrealistic given the current technology for glycoproteomics (as well as glycolipidomics and GAG-"omics"). In the short term the focus should be on worthy targets and improvements of technology to a point where average laboratories can make progress on sequencing, as is true for nucleic acid and protein chemistry. Thus, the available evidence indicates that glycosylation is relatively glycoprotein-specific, site-specific within a glycoprotein, and certainly cell-type specific. In a similar manner, the structures and diversity of proteoglycans/glycosaminoglycans (GAGs) and glycolipids in cells are also relatively cell-type specific.

Sizes of glycan determinants

Since the concept of glycan determinants is inextricably linked to antigens and antibodies, in this review we will consider glycan determinants that are recognized by antibodies and all other types of GBPs. The early work on blood group antigens and other antigens, including dextrans, indicated that antibodies could recognize relatively large determinants. Prof. Kabat and colleagues noted that "These studies demonstrated that the upper limit for the size of an anti-(α1-6)dextran combining site was complementary to six or seven (α1-6)-linked glucoses in dextran, and the lower limit was complementary between one to two such (α1-6)-linked glucoses"[91] [quoted from ref. 92]. These discoveries fit well with studies on blood group antigens in glycolipids, where specific antibodies have been identified that recognize large complex determinants (see Fig. 2 and Table 1 and references therein). The upper limit of recognition for six or seven monosaccharides by a GBP is consistent with studies of other glycan recognizing lectins and enzymes. Lysozyme, which cleaves peptidoglycan and chitin oligosaccharides, binds to a hexasaccharide,[93–95] distorting the chain so that cleavage occurs to release a di- and tetrasaccharide. A variety of *endo*-β-galactosidases also recognize relatively large-sized tetra/hexasaccharides as substrates.[96–98] The tomato lectin binding site was also found to be complex, since the lectin did not bind monosaccharides, and could distinguish large glycan determinants, such as chitin and extended poly-*N*-acetyllactosamine (poly-LacNAc) (-3Galβ1-4GlcNAcβ1-)$_n$, from short glycans with only one Galβ1-4GlcNAcβ-R

determinant.[99–101] Other examples are cited below in the discussion of specific glycan determinants. The available evidence indicates that GBP binding sites may accommodate relatively large glycans, and the typical glycan determinant comprises a ''framework'' from 2 to 6 linear monosaccharides that may contain additional branching modifications.

Many studies have shown that different GBPs recognize complex glycan determinants that range in size from disaccharides recognized by plant lectins to large glyco(sulfo)-peptides recognized by P-selectin, where the sugars and aglycone components are all recognized within the binding site of the GBP (Fig. 1). Despite evidence that GBPs can recognize complex glycan determinants, they are often classified in a simple manner as mannose binding, galactose binding, fucose binding, *etc*., based on whether high concentrations of a simple hapten can inhibit binding. Such classifications are clearly useful, since many lectins have a primary monosaccharide recognition site, and such classifications may promote the applicability of lectins. Inexpensive hapten inhibitors and plant and animal lectins are incredibly useful in isolating glycans and defining sequences of glycoproteins and glycolipid glycans, since binding can often be readily inhibited by high concentrations of inexpensive haptens. Such simple classifications, however, should not be generalized to all GBPs and misdirect us away from the fundamental concept that GBPs typically recognize complex glycan determinants with much higher affinity than displayed toward simple haptens.

Minimal or partial determinants—"necessary but not sufficient"

In the area of glycosciences, the terms minimal or partial determinant are often used to describe the minimal structure of a glycan required for measurable interaction with a GBP.[102–106] However, a minimal determinant may also be considered an element of glycan structure that is ''necessary but not sufficient'' for high affinity interactions. Thus, α-methyl-mannoside is a minimal determinant for the plant lectin ConA, but it is not sufficient in itself to be recognized with high affinity in comparison to high mannose-type *N*-glycan recognition by ConA.[107–109] Mannose-6-phosphate (Man-6-P) is a minimal determinant weakly recognized by P-type lectins in comparison to phosphorylated high mannose-type *N*-glycan recognition.[38,110,111] Similarly, lactose and β-galactosides may be considered as presenting minimal determinants for galectins, but such a minimal determinant is not sufficient for the higher affinity binding observed in the interactions of galectin-3 with blood group A substance or

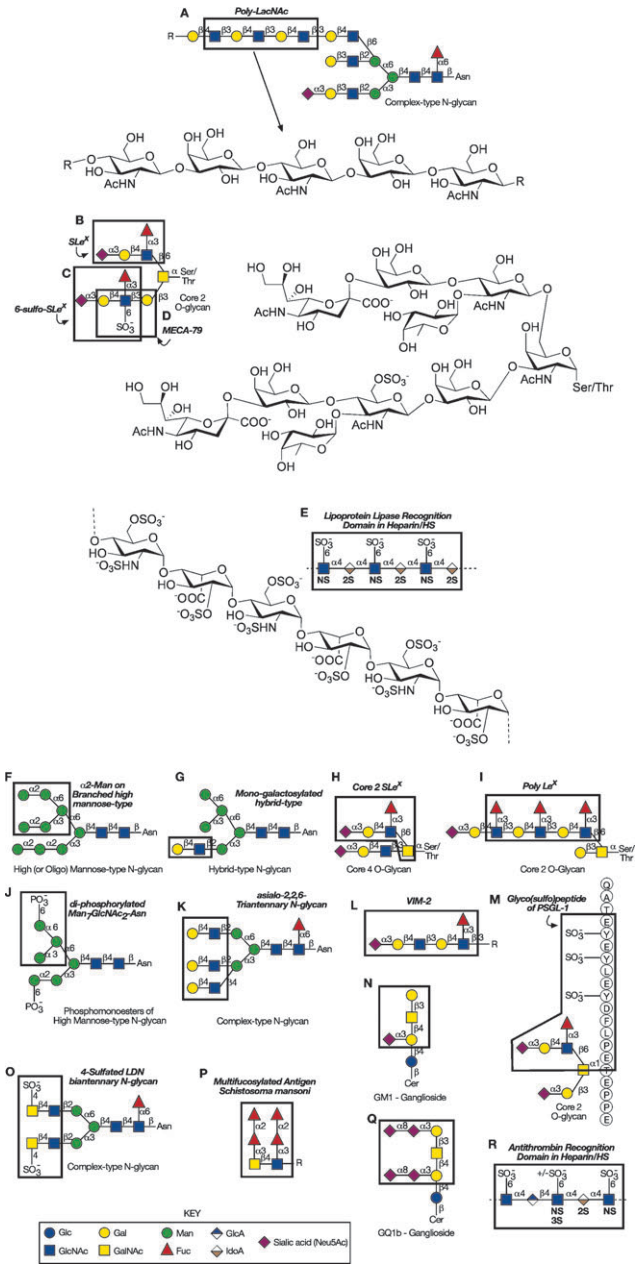


Fig. 1 Structures of complex carbohydrates and examples of glycan determinants within them that are recognized by GBPs. The glycans shown are diverse and comprise *N*- and *O*-glycans, glycosphingolipids and glycosaminoglycans. Symbols used for individual monosaccharides are styled after the convention of the Consortium for Functional Glycomics (www.functionalglycomics.org/), and are indicated in the key. The illustrations of several glycan determinants using ChemBioDraw Ultra® are shown for A–E. The area inside the boxed regions denotes glycan determinants specifically recognized by lectins and/or antibodies (GBPs). In most cases structural isomers are not recognized with equivalent affinity or not detectably bound by the GBPs. A range of glycan determinants are shown from disaccharides to octasaccharides. The boxed determinants are recognized with relatively high affinity and specificity by the following: **A**. galectin-3 and galectin-9;[41,190–195] **B**. mAb CSLEX1 to SLe^x;[213] **C**. L-selectin;[214] **D**. mAb MECA-79 antigen;[215] **E**. heparin/heparan sulfate (H/HS) binding domain for lipoprotein lipase;[216] **F**. concanavalin A (ConA);[107,108,134] **G**. *Ricinus communis* agglutinin-I (RCA-I);[217,218] **H**. mAb CHO-131 to Core 2 SLe^x;[219] **I**. poly-Le^x antibody to circulating cathodic antigen;[220] **J**. P-type lectins;[39] **K**. asialoglycoprotein receptor (Gal/GalNAc lectin) in hepatocytes;[221] **L**. mAb to VIM-2 antigen;[222] **M**. P- and L-selectin;[34,121] **N**. cholera toxin;[183] **O**. cysteine-rich domain of the mannose/GalNAc-4-SO4 receptor;[223] **P**. mAb 114-4D12 to multifucosylated *O*-glycan of *Schistosoma mansoni*;[224] **Q**. anti-GQ1b IgG;[225] **R**. H/HS binding domain for antithrombin.[226–228]

poly-*N*-acetyllactosamine, and the lack of recognition for many compounds containing β-galactosyl residues.[112–114] Minimal determinants for L-selectin binding may be considered to be both GlcNAc-6-sulfate[106] and sialyl Lewis x (SLe^x), whereas E-selectin uses SLe^x as a minimal determinant, and P-selectin uses both tyrosine-sulfate and sialyl Le^x as minimal determinants.[34,115–121] Each selectin demonstrates extreme preference for certain modified SLe^x glycans or glycoproteins that express SLe^x along with other glycan and aglycone determinants,[122,123] and SLe^x itself is only weakly recognized by selectins. Thus, SLe^x is necessary but not sufficient for high affinity interactions of selectins with their physiological ligands. Similarly, Galβ1-3/4GlcNAc is necessary but not sufficient for high affinity interactions of glycans with galectins. Siglec-8 recognizes Neu5Acα2-3Galβ1-R as a minimal determinant, yet binds 6′-sulfo SLe^x but not its isomer 6-sulfo-SLe^x.[124] Influenza virus hemagglutinin, which uses either NeuAca2-3Gal or NeuAca2-6Gal as partial determinants, recognizes specific types of glycans carrying these determinants, thus 2-3 or 2-6-linked sialic acid is necessary but not sufficient for virus binding.[125–127] The concept that partial glycan determinants may be necessary but not sufficient for high affinity interactions recognizes the contribution that additional glycan and aglycone determinants make to GBP recognition, and can contribute to identification of endogenous physiological ligands for GBPs, as discussed below.

How many glycan determinants exist?

The total number of glycan determinants in the human glycome can be estimated based on a variety of published observations on glycan structures and their recognition by GBPs (lectins, adhesins, antibodies, toxins, *etc*.) and a set of assumptions that derive from this structural information. The assumptions are that (a) each determinant is composed of a "framework" sequence containing two to six *linear* monosaccharides; (b) the complex glycan determinants comprise a variety of partial determinants linked together; (c) selected linkages are used based on known linkage motifs and anomeric configurations; and (d) glycan modifications are to be considered, such as sulfation and phosphorylation. Since phosphorylation is an atypical modification found primarily on *N*-glycans of lysosomal enzymes, only a select number of phosphorylated mannose glycans are considered here. Some of the typical framework sequences include -Galβ-GlcNAcβ-Manα1-3/6-, –GalNAcβ-GlcNAcβ-Manα1-3/6-, -Galβ-GlcNAcβ-Galβ1-, *etc*. Such sequences are common in animal glycoproteins. We should also consider the partial determinants, which are modified residues along with differential linkages within the framework sequence. These include Neu5Acα2-6Galβ1-3, Neu5Acα2-6Galβ1-4, 9-*O*-acetyl-Neu5Acα2-6Galβ1-3, 9-*O*-acetyl-Neu5Acα2-6Galβ1-4, Neu5Acα2-3Galβ1-3, Fucα1-2Galβ1-4, and other modified residues within the framework -Galβ-GlcNAcβ-Manα1-3/6-, including GlcNAc modifications, such as GlcNAcβ2, GlcNAcβ3, GlcNAcβ4, GlcNAcβ6, 6-*O*-sulfo-GlcNAcβ2, 6-*O*-sulfo-GlcNAcβ3, and Man linkages or modifications, such as Manα3, Manα6, and Manβ1-4, *etc*. Many of these glycan determinants are shared in animal cell *N*- and *O*-glycans and glycolipids. Some linkages, such as extended glycans on "bisected" *N*-glycans, including Galβ1-4GlcNAcβ1-4Manβ1-4GlcNAc-R and other modifications,[128,129] may not be common, but their recent discovery suggests the presence of a large possible repertoire of such sequences. Lists of these partial determinants are shown in Fig. 2. When these partial determinants are linked in multiple possible combinations, a total of 1722 structures can occur, which includes the phosphomannose-type modifications. This is a conservative number, since it does not include many other potential partial determinants of Gal, including the dozens of potential sialic acid modifications[130] or polysialic acids,[131,132] which may represent independent glycan determinants. For example, a murine IgG2a binds polysialic acid with a binding site that accommodates at least eight sialic acid residues.[133] The number of 1722 glycan determinants also does not reflect the combinatorial properties of determinants in branched glycans. For example, branched *N*-glycans could have different types of glycan determinants on each branch and together these might constitute a new glycan determinant, although it would have to be functionally shown that such combined determinants are recognized by a GBP. It is surprising that this property of GBPs has not been studied in greater detail, considering it is clearly the case for the common plant lectin ConA, which binds branched mannose Manα1-3(Manα1-6)Manα1-R,[107,108,134] and L-PHA, which only binds complex-type *N*-glycans with 2,6-branched mannose Galβ1-4GlcNAcβ1-2(Galβ1-4GlcNAcβ1-6)Manα1-6R,[135,136] the anti-Ii blood group antibodies,[137] as well as the 2G12 mAb against HIV high mannose-type *N*-glycans that recognize $Man_9GlcNAc_9$.[36]

Additionally, a larger number of other glycan modifications or motifs have been found in animals, and they may occur to some extent in the human and murine glycomes, but they were not included in these calculations for Fig. 2. For example, Glc-terminated glycans, which were earlier thought to not occur significantly in animal glycomes, are found in several types of glycans, such as Glcα1-3Manα1-2Manα-R, and in fact are key intermediates in *N*-glycan biosynthesis and glycoprotein quality control.[138,139] These and other motifs are beyond the scope of this review, but certainly represent novel linkages and modifications, including Glcβ1-3Fucα1-Ser/Thr in EGF modules,[140] Siaα2-3Galβ1-4GlcNAcβ1-3Fucα1-Ser/Thr in notch,[141] and Glcα1-2Galβ1-hydroxylysine in collagen[142] and the serum mannose binding lectin.[143,144] In addition to the well known phosphorylation of mannose to give mannose-6-phosphate, other glycans may also be phosphorylated and sulfated, as seen in -GlcUAβ1-3Gal(±6-*O*-sulfate)-β1-3Gal(±6-*O*-sulfate)β1-4Xyl and -GlcUA-Gal-Gal-Xyl-(2-*O*-phosphate)- of the linkage region of proteoglycans.[145,146] Other unusual structures, such as Glcβ1-6(Galβ1-3)GalNAcα1-Ser/Thr and Glcβ1-6(Glcβ1-3)GalNAcα1-Ser/Thr,[147] have been identified in *Caenorhabditis elegans* where they are abundant, and raise questions whether they might also occur in higher animals. Thus, the list of partial determinants identified here, and even the framework structures, are likely to be an underestimate, but this approach allows one to readily categorize and expand the list of theoretical glycan determinants as new glycan structural information arises.

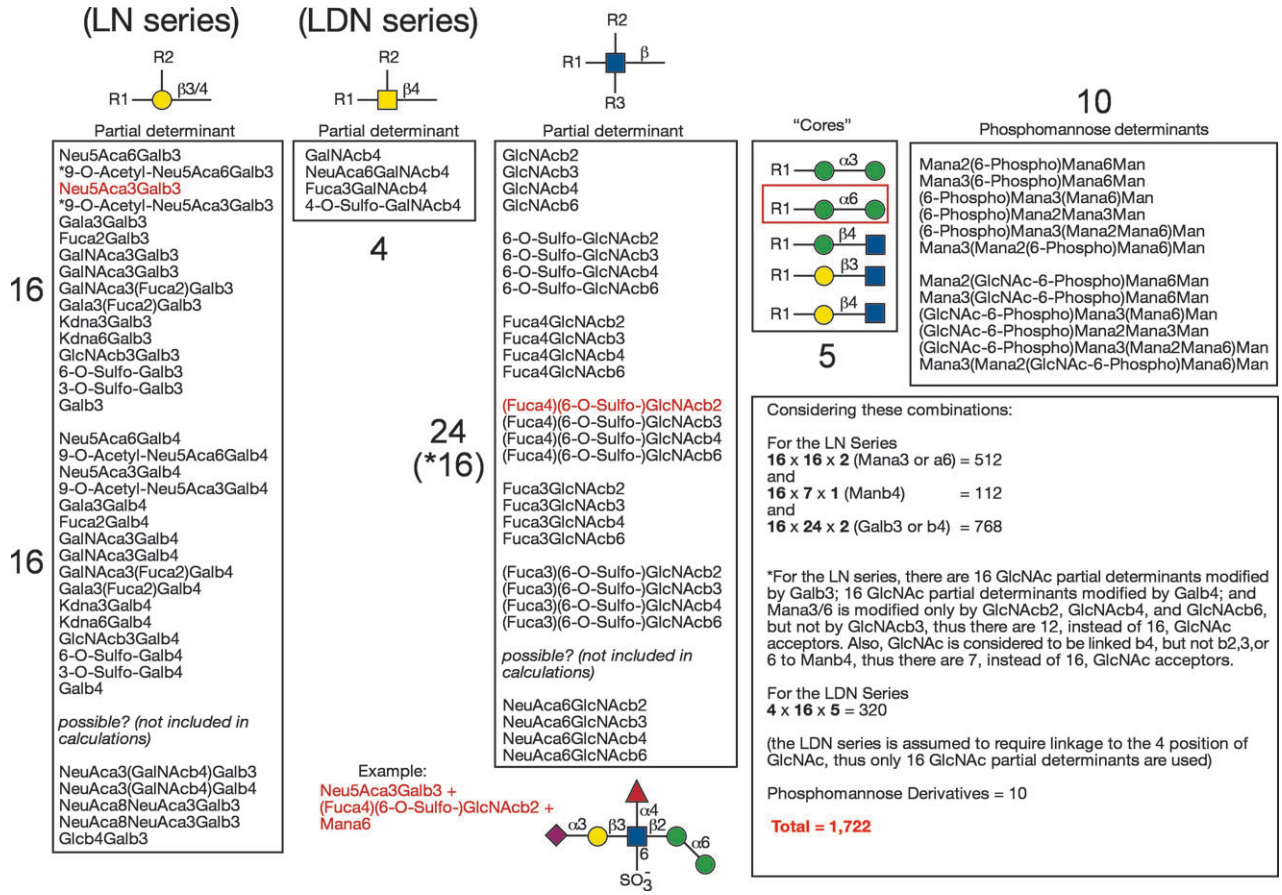


Fig. 2 The assembly of partial determinants into specific glycan determinants, as found in animal cell *N*- and *O*-glycans and glycolipids. The partial determinants are envisioned as derivatives of either galactose (Gal) or *N*-acetylgalactosamine (GalNAc), which are linked to derivatives of GlcNAc, to create Gal/GalNAc-GlcNAc disaccharides, such as *N*-acetyllactosamine (LN) or LacdiNAc (LDN). These LN or LDN disaccharides may be linked to a third monosaccharide, either mannose (Man) or Gal, to generate a trisaccharide backbone. For example, multiplying the number of determinants in the Gal derivatives by those in the GlcNAc derivatives, and by those in either the Man or Gal derivatives provides for the possible number of glycan determinants as shown in the inset box, which for this consideration with both LN and LDN gives 1722 possible glycan determinants. Also, for the case of branched *N*-glycans, another factor to consider is that one branch may contain one type of glycan determinant and the other could contain an entirely different determinant, but such combined determinants are not considered here for calculations. In addition, other possible extensions of glycans, polysialic acid, and other modifications of Gal partial determinants are not used in the calculations. Also, combinations of Gal and GlcNAc partial determinants occur to generate poly-*N*-acetylactosamine (poly-LacNAc) chains, but these are also not included in the calculations. For simplicity, a and b are used in place of alpha and beta for the linkages. The red lettered partial determinants are linked together to show an example of how the glycan determinants are constructed.

Additional frameworks exist for *O*-glycan core structures based on GalNAcα1-Ser/Thr, as shown in Fig. 3. GlcNAc and Gal residues within these *O*-glycan frameworks can be modified and extended in a large variety of ways to generate at least 928 combinations. This does not include the *O*-glycans based on cores 5, 7, and 8 for which little information exists. Also, these 928 combinations do not take into account the additional core modifications, as indicated by the dotted arrows in the cores in Fig. 3, or modifications of sialic acids. In both Fig. 2 and 3, the partial determinants can be added together to give even larger glycan determinants than those indicated. For example, poly-*N*-acetyllactosamine and poly-fucosylated poly-*N*-acetyllactosamine (type 2 chains) could be constructed by linking Gal-GlcNAc-Gal-GlcNAc, *etc*., to create large framework glycans containing penta-, hexa-, hepatasaccharides, *etc*. Unusual *O*-glycan modifications are being found increasingly through glycomic analyses of normal and diseased tissues. For example, there is an increased expression of a core 3 SLe^x hexasaccharide NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3(NeuAcα2-6)GalNAcα1-Ser/Thr in colon cancer cells and a decreased expression of its sulfo-Le^x counterpart, sulfate-3Galβ1-4(Fucα1-3)GlcNAcβ1-3(NeuAcα2-6)-GalNAcα1-Ser/Thr, in colon cancer cells.[148]

The repertoire of glycolipids (glycosphingolipids and glycophospholipids) in animal glycomes is not well understood, but there are at least ~400 different types of glyco(sphingo and glycero)lipids, based on the list of glycans at the Lipid Maps Consortium website (www.lipidmaps.org/). Most mammalian glycolipids are comprised of neutral and acidic species of glycosphingolipids that contain common core structures (Fig. 4). It is not possible to generate a straightforward combinatorial approach using partial determinants, to generate the known and predicted glycolipid glycan determinants, as in Fig. 2 and 3, since the partial determinants are not well

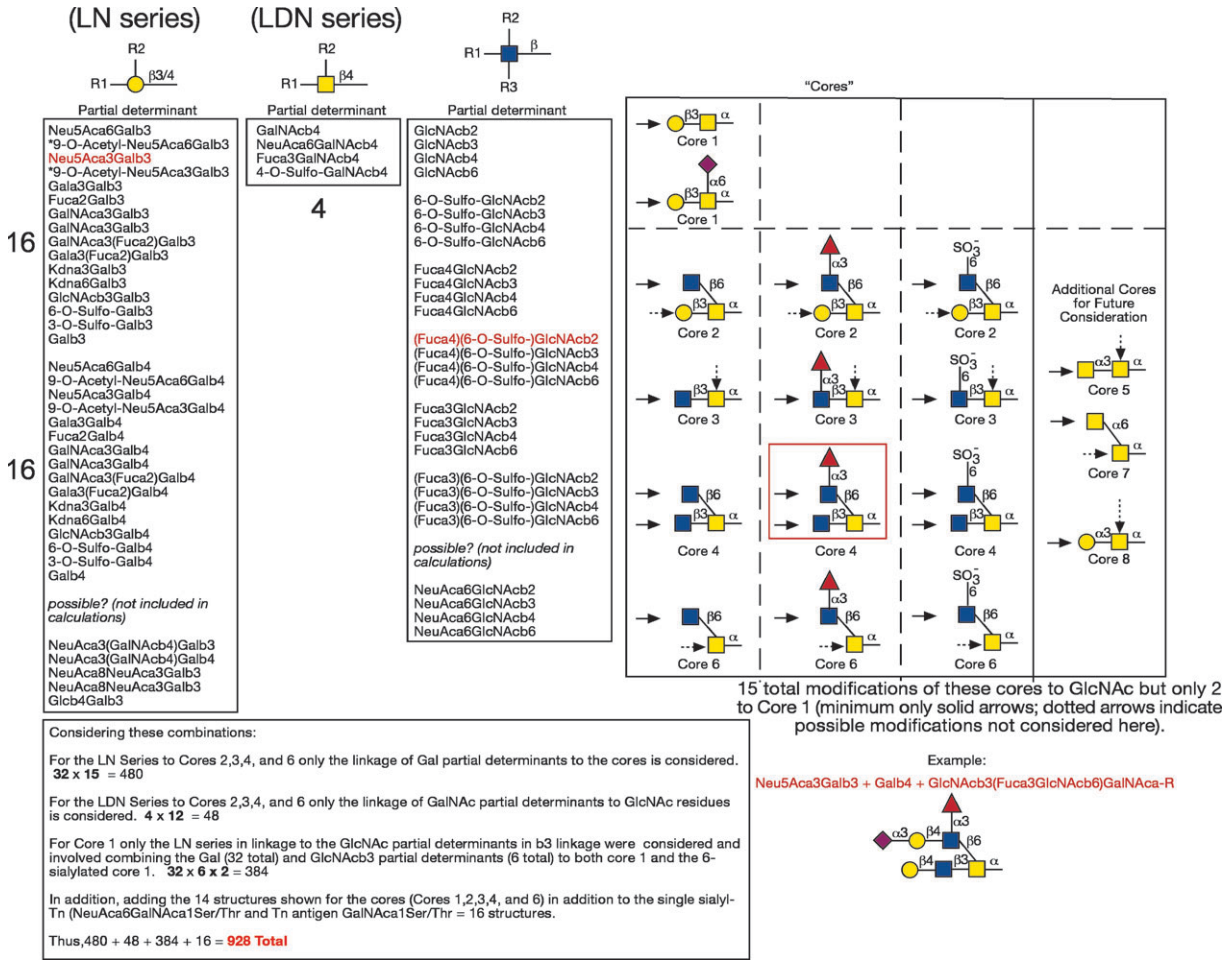


Fig. 3 The assembly of partial determinants into specific *O*-glycan determinants found in Ser/Thr-linked mucin-type structures. As for Fig. 2, the partial determinants are envisioned as derivatives of either galactose (Gal) or *N*-acetylgalactosamine (GalNAc), which are linked to derivatives of GlcNAc, to create Gal/GalNAc-GlcNAc disaccharides, such as *N*-acetyllactosamine (LN) or LacdiNAc (LDN). These LN or LDN disaccharides may be linked to a GlcNAc residue in one of the five positions shown in the solid arrow for the 4 types of core *O*-glycans shown, cores 1, 2, 3, 4, and 6. The dotted arrow indicates at least one other potential linkage, but that is not used in the calculations. Also, the other core *O*-glycan linkages are shown, cores 5, 7, and 8, but they are not considered in the calculations shown in the inset box, to give 928 possible glycan determinants. As discussed in Fig. 2, additional modifications and extensions of glycans with other Gal and GlcNAc partial determinants are possible. For simplicity, a and b are used in place of alpha and beta for the linkages. The red lettered and boxed partial determinants are linked together to show an example of how the glycan determinants are constructed.

understood. A list of common glycolipids and glycolipid antigens is shown in Table 1, which is provided to show that many of these glycolipid glycans are themselves complete determinants that are uniquely recognized by antibodies and other GBPs. Selected references to studies identifying mAbs or other GBPs that recognize many of these glycolipid are provided there. Remarkably, many of these glycolipids have been identified from earlier studies beginning in the 1970's by groups led by Prof. Hakomori, Prof. Feizi, Prof. Kaproski, and Prof. Ginsburg, and others cited in Table 1, who developed mAbs that first demonstrated the remarkable specificity of antibodies for large glycan determinants. While *N*- and *O*-glycans can share glycan determinants with glycolipids, there is considerable evidence that many glycolipid structures, such as those in the *ganglio* series, are not shared, and many glycolipids are specifically recognized by antibodies that do not recognize glycoproteins.[149,150] In addition, the glycosylphosphatidylinositol (GPI) anchor structure of GPI-anchored membrane glycoproteins, which can occur as free glycophospholipid structures, are included. The common mammalian GPI contains the core structure ethanolamine-PO_4-6-Manα1-2Manα1-6Manα1-4GlcNH_2-α1-6-*myo*-inositol-1-PO_4-lipid, but there are a variety of modifications that can occur on each mannose residue and the inositol residue[151–153] (Fig. 4). While the number of such combinations in humans is not well understood, it could be up to at least one hundred. Such glycophospholipid determinants are differentially recognized by antibodies,[154,155] as has also been shown recently for serum antibodies in patients with malaria using microarray analyses of synthetic GPI glycoconjugates.[156] In summary, while the total number of glycolipid determinants in glycosphingolipids and glycophospholipids is difficult to estimate, for the considerations here we will assume that the glycolipid repertoire of glycan determinants is $\sim$500.

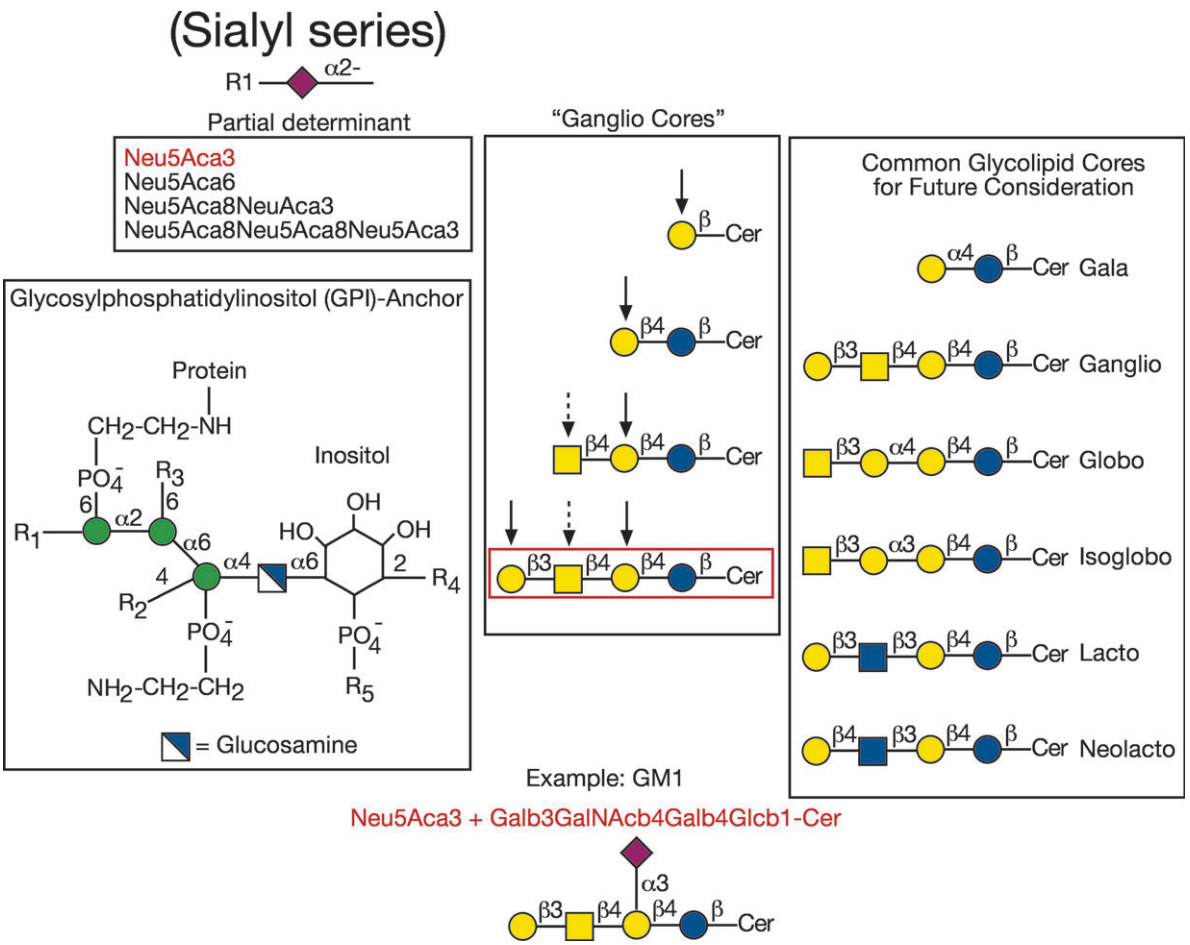


Fig. 4 The partial determinants of glycosphingolipids. Only the simple *ganglio* series is shown here as being modified by sialic acid partial determinants, which are typically Neu5Ac rather than Neu5Gc. The solid arrow indicates sites known to be modified by Neu5Acα2-3, whereas the dotted arrow indicates the site modified by Neu5Acα2-6. The red lettered and boxed partial determinants are linked together to show an example of how the glycan determinants are constructed. For simplicity, a and b are used in place of alpha and beta for the linkages in the written structures. It is possible that Neu5Gc could also occur, but most studies have shown that brain gangliosides from all mammals contain Neu5Ac rather than Neu5Gc.[229] Other known glycolipid core structures, in addition to the *ganglio* series, are shown to the right. Other sugar residues and sulfate can be added to all of these core sequences. A larger list of known animal gangliosides and other glycolipids is shown in Table 1. On the lower left side is a typical core structure for human glycosylphosphoinositol (GPI) anchored glycoproteins. The R_{1-5} refer to sites of modification by additional moieties, such as R_1 = Man or OH, R_2 = GalNAc or OH, R_3 = ethanolamine phosphate or OH, R_4 = palmitate or OH; and R_5 = alkyl or diacyl glycerol.

The potential diversity of glycosaminoglycans (GAGs) is more straightforward to predict, given the known modifications of each residue within hyaluronan, keratan sulfate, chondroitin/dermatan sulfates, heparin/heparan sulfates.[157] Consider that GAG recognition by GAG-binding proteins can require up to 5 monosaccharide units, and a variety of microarray applications have been explored on this principle.[158–167] Thus, for defining the glycan determinants in GAGs, it is useful to consider all possible pentasaccharides. The "flanking" sugar residues at either end of a pentasaccharide can alter or contribute to its recognition, and clearly GAGs beyond pentasaccharides can be recognized.[157,167,168] There are additional GAG modifications that are not included in this calculation, such as the findings of Prof. Kannagi's group in regard to antibodies to heparan sulfate determinants,[169] which include acharan sulfate $(IdoA2S\text{-}GlcNAc)_n$, acharan $(IdoA\text{-}GlcNAc)_n$ and *N*-acetyl-heparosan $(GlcA\text{-}GlcNAc)_n$. The partial determinants for GAG recognition include GlcNAc, GalNAc, GlcA, and IdoA, and their *N*- and *O*-sulfated versions, as shown in Fig. 5. Given the combinations of each partial determinant, it is possible to generate 3981 pentasaccharides (~4000).

Thus, considering the total possible combinations of partial determinants in Fig. 2–4 there are ~3000 glycan determinants in *N*- and *O*-glycans and glycolipids, along with the ~4000 possible GAG pentasaccharides. Thus, there are possibly ~7000 glycan determinants in the human glycome. This number is probably an underestimate, since a conservative number of combinations are used and combinations of these glycan determinants in branched configurations would greatly expand the repertoire, if they, such combined glycan determinants, are found to be recognized by a specific GBP. In addition, the sites of attachment of glycans to proteins and lipids and adjacent aglycone moieties can contribute or alter

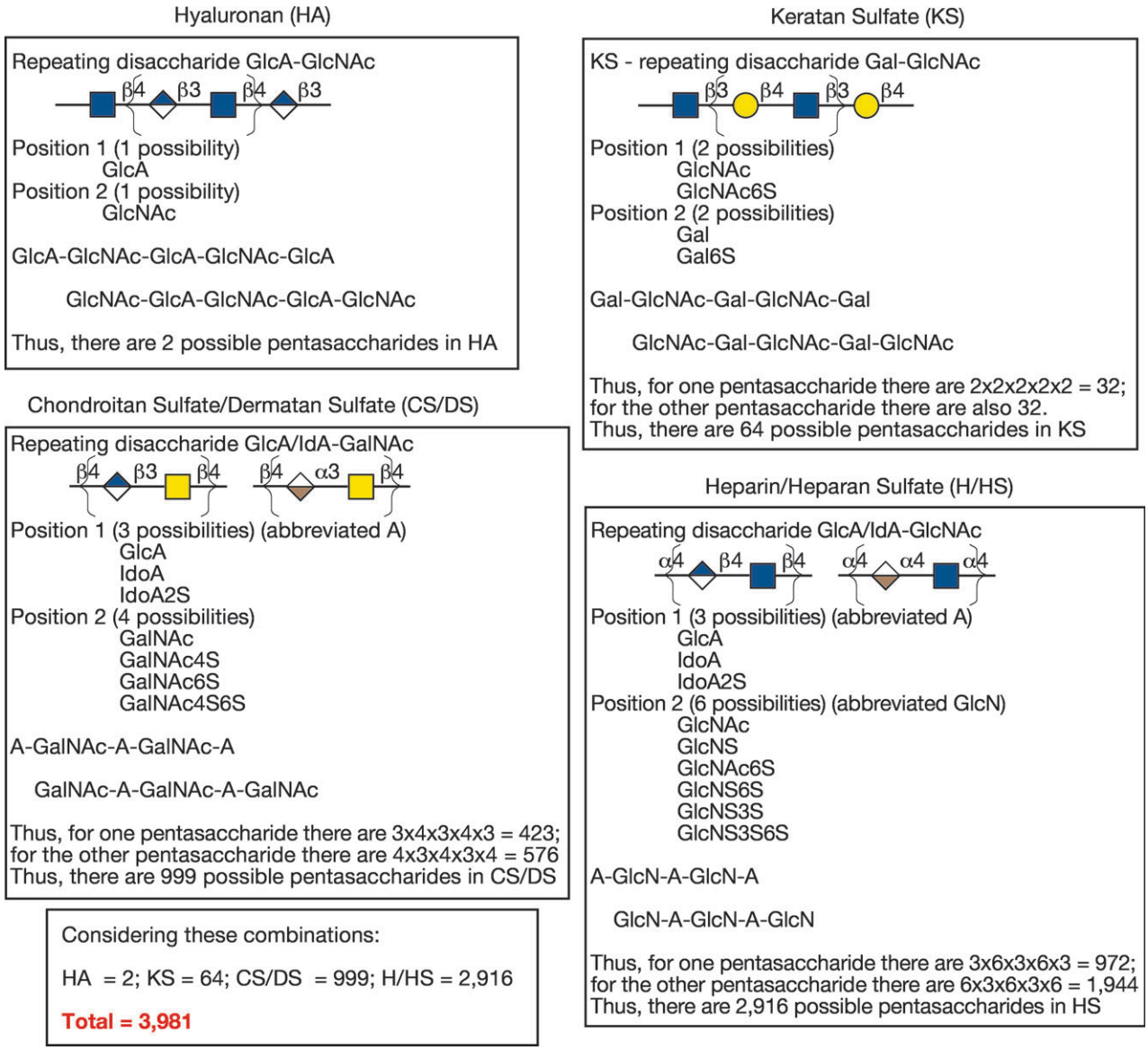


Fig. 5 The assembly of partial determinants within known glycosaminoglycans to generate specific pentasaccharide glycan determinants. Each type of glycosaminoglycan is considered as a pentasaccharide, and since the GAGs are repeating disaccharides, there would be at least two types of pentasaccharides for each, as indicated. All known modifications of Gal, GlcNAc, GalNAc, GlcA, and IdoA are shown and utilized in all possible combinations. The total number of H/HS pentasaccharides is, as expected, the largest set (total = 2916), while the hyaluronan (HA) pentasaccharides are the smallest set (total = 2). Altogether, there are 3981 possible glycosaminoglycan pentasaccharide combinations.

glycan recognition by GBPs. Many levels of combinations of glycan and aglycone determinants, in addition to multivalent presentations,[23,170] contribute to the essential functions of the glycan determinants.

Low *versus* high affinity binding of glycans by GBPs

The affinity of GBPs for monosaccharides and small glycans (di- and trisaccharides) is relatively low (K_d = 10^{-3} to 10^{-6} M), but binding to larger glycan determinants is typically much higher (K_d = 10^{-6} to 10^{-8} M). In addition, affinity between multivalent lectins and multivalent ligands can lead to cross-linking and stable lattice formation.[171–173] It should be remembered that a particular glycan often is found in abundance at the cell surface or on a particular glycoconjugate, and thus, while binding affinities might be considered low to isolated glycans, the physiological apparent affinity (avidity) of a GBP to the multivalent presentation might enhance affinity by orders of magnitude. For example, it has been estimated that the concentration of sialic acid at the surface of a B lymphocyte may be above 100 mM.[174] While not well studied in regard to multimeric animal lectins, multiple glycan-binding domains within a single GBP can lead to higher affinity (avidity) interactions. The situation is analogous to that of antibodies, where the divalent affinity of an antibody is typically only 10–20 times higher than the monovalent affinity.[175–178]

A good example of biological multivalency is the recognition of phosphorylated lysosomal hydrolases containing Man-6-P determinants by the membrane-bound P-type lectins

(CI-MPR and CD-MPR).[39] The CI-MPR (cation-independent MPR) contains multiple domains that bind Man-6-P, while the CD-MPR (cation-dependent MPR) is dimeric with each monomer having a single binding site. While both MPRs bind the monosaccharide Man-6-P with equivalent affinity ($K_d \approx 7–8 \times 10^{-6}$ M), they bind much better to the extended Man-containing glycan having Man-6-P residues and recognize at least an α1,2-linked phosphorylated trimannoside.[38] In addition, the multivalent interactions between the receptors and their phosphorylated glycans give an avidity in the range of ($K_d \approx 1–10 \times 10^{-9}$ M) for both MPRs. Importantly, domain 5 of the CI-MPR selectively binds GlcNAc-P-6-Man-R phosphodiesters.[179]

Direct monovalent binding to glycan determinants can also be of relatively high affinity. One of the best studied examples of this is cholera toxin, which specifically recognizes GM1 ganglioside with high affinity, and binds to a specific tetrasaccharide determinant within the ganglioside GM1 (Fig. 1).[180–182] Notably, the individual B subunit of cholera toxin binds GM1 oligosaccharide in solution in the nanomolar range, with a relatively high affinity ($K_d \approx 40 \times 10^{-9}$ M),[183] and may exhibit some slight positive cooperativity.[184] Native cholera toxin with five B subunits and one A subunit (A_1B_5) binds ganglioside GM1 on cells and liposomes with even higher affinity, ranging from $K_d \approx 4.6 \times 10^{-10}$ M[185] to $K_d \approx 4.6 \times 10^{-12}$ M.[186] Clearly, as discussed previously,[183] cholera toxin recognizes the composite glycan with a specific and high affinity interaction that is more than the sum of the individual, weak binding monosaccharide determinants.

Another example of high affinity monovalent binding to glycan determinants is seen in galectins. Galectin-1 was the first animal lectin to be crystallized[187,188] and it was co-crystallized with lactose and *N*-acetyllactosamine. While the name galectin was chosen for this class of lectins that all seem to recognize β-galactosides,[189] each galectin demonstrates extreme preference for different glycan determinants. For example, both galectin-3 and -9 recognize complex glycans with much higher affinity and specificity than simple glycans.[41,190–193] Human galectin-3 binds lactose with a $K_d \approx 55 \times 10^{-6}$ M, but binds to poly-*N*-acetyllactosamine (poly-LacNAc) with $K_d \approx 3 \times 10^{-6}$ M, independently of the terminal residues on the poly-LacNAc.[192] These results are consistent with much earlier studies showing a preference of galectin-3 for poly-LacNAc-containing glycans.[112] Such glycan determinants were also readily observed to be specifically recognized compared to isomeric glycans in glycan microarray analyses of human galectin-3.[194] Similarly, galectin-9 binds short *N*-acetyllactosamine chains with low affinity but binds to longer poly-LacNAc sequences with $K_d \approx 9 \times 10^{-8}$ M.[190] Co-crystals of galectin-9 with poly-LacNAc-containing glycans show specific recognition of the three continuous residues -GlcNAcβ1-3Galβ1-4GlcNAcβ1- within poly-LacNAc.[41] These and other data on the cellular ligands for these galectins[191,192,195–198] indicate that poly-LacNAc-containing glycans represent the glycan determinants recognized by these galectins. The *C. elegans* galectin LEC-6 binds weakly to *N*-acetyllactosamine, but with much higher affinity ($K_d = 14.7 \times 10^{-6}$ M) to the unusual trisaccharide determinant Gal-Fuc-GlcNAc, within the context of an *N*-glycan with the structure Hex-methylFuc-Man-(Man)-Man-GlcNAc-(Gal-Fuc-)-(Fuc)GlcNAc.[199] Thus, each galectin probably has a unique glycan determinant that is recognized with high affinity.

Relationship of glycan determinants to the physiological ligands for GBPs

While an understanding of the nature of glycan determinants and modes of GBP binding to such determinants is important, this knowledge is still far short of defining the physiological ligand for the GBP. Identifying the physiological ligands for GBPs is difficult in most cases, largely because it may require a large body of knowledge of the GBP in terms of its sites and level of expression, valency, and information about potential targets to which it binds, along with a combination of skills in cell biology, chemistry, biochemistry, and molecular biology, *etc*. Approaches that are useful in functional proteomics, as yeast 2-hybrid or tandem affinity purification (TAP) methodology, are not possible in studies of glycan interactions with GBPs. Historically, the most common techniques for defining physiological ligands for GBPs involved affinity chromatography and pull-down affinity techniques, or genetic modifications of glycosylation pathways in cell lines, animals, and humans. Defining glycan determinants for GBPs using glycan microarray technologies is an appropriate new tool for such studies to explore the possible physiological ligands for a GBP, but the diversity of the glycans on existing microarrays need to be enlarged to match the predicted diversity of glycan determinants, and consideration must be given to the fact that glycans are typically linked to an aglycone moiety, *e.g.* proteins or lipids. So far, the Consortium for Functional Glycomics has made great headway in this quest to generate a large diversity of glycan determinants (outside of glycosaminoglycans), using chemical and/or enzymatic approaches. This has resulted in a microarray containing nearly 500 compounds, and a majority of them represent glycan determinants important for high affinity and specific glycan binding by GBPs.[58,170,200–203] This and other microarrays[48,194,204–208] have been incredibly useful in defining glycan recognition by GBPs, and they provide a valuable tool in the quest to identify the physiological ligands and glycan determinants for GBPs. In addition, there are many potential strategies for generating glycan microarrays and immobilizing glycans on solid surfaces to enable studies on GBP recognition.[209,210] The ongoing development of glycosaminoglycan microarrays[211,212] will also be useful in the future, and as these arrays grow their diversity will undoubtedly reflect the pentasaccharide diversity of the GAGs. While the identification of glycans recognized by GBPs using microarray technologies is useful, the results should be cautiously interpreted in terms of whether they provide clues to the physiological ligands. Obviously, microarrays can lead to testable biological predictions and such studies should be carried out in order to define the physiological ligand(s) for a given GBP.

Summary

Fischer's concept of the lock-and-key recognition of carbohydrates has come a long way in the past decades and we are gaining a better appreciation of the mystery of why there are

so many types of glycans and structural modifications of glycoconjugates. As a result, the field of glycoscience research is undergoing a tremendous growth in prestige and understanding. The identification of glycan determinants recognized by GBPs is a major goal of glycoscience research, and will lead to a better functional appreciation of human and animal glycomes.

While the complexity and diversity of the glycome in terms of total glycan structures is not calculable, the number of glycan determinants likely to be important in GBP recognition is an estimable number. The list of glycan determinants described here is likely to grow and diversify as more glycan structures are completely elucidated and their interactions with GBPs are explored. The number of glycan determinants estimated here at $\sim$7000 is within the scope of chemical and enzymatic synthesis over the next decade, and the production of these glycans should be a major goal of the glycoscience community. The availability of glycan microarrays to identify glycan determinants has helped fuel a revolution in the glycosciences, since only minor amounts of glycans are needed, and correspondingly minor amounts of potential GBPs. Thus, glycan synthesis can be appropriately scaled down to allow practical production of highly diverse glycan libraries of glycans. In addition, having these determinants will allow production of antibodies against them, for example, by using conventional mouse monoclonal strategies or phage display. Importantly, glycan microarrays would help to identify antibody specificities and allow the use of well-defined antibodies in the field. Such antibodies also would be invaluable tools in glycobiology to identify glycans and glycoconjugates expressing these determinants, both *in situ* and in isolated materials.

Acknowledgements

This work was supported by grants from the National Institutes of Health (HL085607), (GM085448), and by resources from the Consortium for Functional Glycomics (Core H), funded by NIGMS/NIH (GM62116). Special thanks are given to Prof. David F. Smith, Dr Jamie Heimburg-Molarino and Ms. Sandra Cummings for helpful discussions and reading the manuscript, and to Prof. Xuezheng Song for help in preparing the illustrations.

References

1 Einfluss der Configuration auf die wirkung der Enzyme, *Ber. Dtsch. Chem. Ges.*, 1894, **27**, 2985–2993; translated by B. Homstedt and G. Liljestrand (ed.), *Readings in Pharmacology*, 1963, pp. 251–3993.
2 , Croonian Lecture, On Immunity with Special Reference to Cell Life, *Proc. R. Soc. London*, 1900, **66**, 433–434.
3 P. Ehrlich, in *Nobel Lectures: Physiology or Medicine*, ed. S. b. t. N. Foundation, Elsevier Publishing Co., Amsterdam, 1967, vol. 304, pp. 1901–1921.
4 F. A. Troy 2nd. in *Biology of the Sialic Acids*, ed. A. Rosenberg, Plenum Press, New York, 1994, pp. 95–144.
5 K. Landsteiner, *Wien. Klin. Wochenschr.*, 1901, **14**, 1132–1134.
6 W. M. Watkins and W. T. Morgan, *Nature*, 1952, **169**, 825–826.
7 W. T. Morgan and W. M. Watkins, *Glycoconjugate J.*, 2000, **17**, 501–530.
8 E. A. Kabat, *Blood Group Substances. Their Chemistry and Immunochemistry*, Academic Press Inc., New York, 1956.
9 N. Sharon and H. Lis, *Adv. Exp. Med. Biol.*, 2001, **491**, 1–16.
10 N. Sharon, *Biochem. Soc. Trans.*, 2008, **36**, 1457–1460.
11 I. J. Goldstein and C. E. Hayes, *Adv. Carbohydr. Chem. Biochem.*, 1978, **35**, 127–340.
12 I. J. Goldstein, *J. Agric. Food Chem.*, 2002, **50**, 6583–6585.
13 H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook and C. Zardecki, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2002, **58**, 899–907.
14 A. Garcia-Pino, L. Buts, L. Wyns, A. Imberty and R. Loris, *Plant Physiol.*, 2007, **144**, 1733–1741.
15 C. S. Wright, *Curr. Opin. Struct. Biol.*, 1997, **7**, 631–636.
16 W. I. Weis and K. Drickamer, *Annu. Rev. Biochem.*, 1996, **65**, 441–473.
17 J. M. Rini and Y. D. Lobsanov, *Curr. Opin. Struct. Biol.*, 1999, **9**, 578–584.
18 J. M. Rini, *Annu. Rev. Biophys. Biomol. Struct.*, 1995, **24**, 551–577.
19 T. Morita, *Pathophysiol. Haemostasis Thromb.*, 2005, **34**, 156–159.
20 R. U. Lemieux, L. T. Delbaere, H. Beierbeck and U. Spohr, *Ciba Found. Symp.*, 1991, **158**, 231–245; discussion 245–238.
21 R. Loris, *Biochim. Biophys. Acta*, 2002, **1572**, 198–208.
22 D. C. Kilpatrick, *Biochim. Biophys. Acta*, 2002, **1572**, 187–197.
23 H. J. Gabius, *Biochem. Soc. Trans.*, 2008, **36**, 1491–1496.
24 A. Imberty and A. Varrot, *Curr. Opin. Struct. Biol.*, 2008, **18**, 567–576.
25 A. Meyer, W. Rypniewski, M. Szymanski, W. Voelter, J. Barciszewski and C. Betzel, *Biochim. Biophys. Acta*, 2008, **1784**, 1590–1595.
26 K. A. Kulkarni, S. Katiyar, A. Surolia, M. Vijayan and K. Suguna, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2008, **64**, 730–737.
27 J. J. Ding, J. K. Bao, D. Y. Zhu, Y. Zhang and D. C. Wang, *Protein Pept. Lett.*, 2008, **15**, 411–414.
28 J. L. Meagher, H. C. Winter, P. Ezell, I. J. Goldstein and J. A. Stuckey, *Glycobiology*, 2005, **15**, 1033–1042.
29 A. B. Boraston, B. S. McLean, J. M. Kormos, M. Alam, N. R. Gilkes, C. A. Haynes, P. Tomme, D. G. Kilburn and R. A. J. Warren, in *Recent Advances in Carbohydrate Bioengineering*, ed. H. J. Gilbert, G. J. Davies, B. Henrissat and B. Svensson, Royal Society of Chemistry, Cambridge, UK, 1999, pp. 202–211.
30 A. Varki, M. Etzler, R. D. Cummings and J. D. Esko, in *Essentials of Glycobiology*, ed. A. Varki, R. D. Cummings, J. D. Esko, H. Freeze, P. Stanley, C. Bertozzi, G. W. Hart and M. Etzler, Cold Spring Harbor Laboratory Press, Boston, 2009, pp. 375–386.
31 *Animal Lectins: A Functional View*, ed. G. R. Vasta and H. Ahmed, CRC press, Boca Raton, FL, 2009.
32 H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163–2166.
33 L. Chatwell, A. Holla, B. B. Kaufer and A. Skerra, *Mol. Immunol.*, 2008, **45**, 1981–1994.
34 W. S. Somers, J. Tang, G. D. Shaw and R. T. Camphausen, *Cell*, 2000, **103**, 467–479.
35 H. Attrill, A. Imamura, R. S. Sharma, M. Kiso, P. R. Crocker and D. M. van Aalten, *J. Biol. Chem.*, 2006, **281**, 32774–32783.
36 D. A. Calarese, C. N. Scanlan, M. B. Zwick, S. Deechongkit, Y. Mimura, R. Kunert, P. Zhu, M. R. Wormald, R. L. Stanfield, K. H. Roux, J. W. Kelly, P. M. Rudd, R. A. Dwek, H. Katinger, D. R. Burton and I. A. Wilson, *Science*, 2003, **300**, 2065–2071.
37 D. A. Calarese, H. K. Lee, C. Y. Huang, M. D. Best, R. D. Astronomo, R. L. Stanfield, H. Katinger, D. R. Burton, C. H. Wong and I. A. Wilson, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13372–13377.
38 L. J. Olson, O. Hindsgaul, N. M. Dahms and J. J. Kim, *J. Biol. Chem.*, 2008, **283**, 10124–10134.
39 N. M. Dahms, L. J. Olson and J. J. Kim, *Glycobiology*, 2008, **18**, 664–678.
40 B. W. Han, B. R. Herrin, M. D. Cooper and I. A. Wilson, *Science*, 2008, **321**, 1834–1837.
41 M. Nagae, N. Nishi, T. Murata, T. Usui, T. Nakamura, S. Wakatsuki and R. Kato, *Glycobiology*, 2009, **19**, 112–117.

42 N. Shah, D. A. Kuntz and D. R. Rose, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 9570–9575.
43 K. Singh, A. G. Gittis, P. Nguyen, D. C. Gowda, L. H. Miller and D. N. Garboczi, *Nat. Struct. Mol. Biol.*, 2008, **15**, 932–938.
44 Z. Li, M. Kienetz, M. M. Cherney, M. N. James and D. Bromme, *J. Mol. Biol.*, 2008, **383**, 78–91.
45 G. J. Swaminathan, D. G. Myszka, P. S. Katsamba, L. E. Ohnuki, G. J. Gleich and K. R. Acharya, *Biochemistry*, 2005, **44**, 14152–14158.
46 H. J. Gabius, H. C. Siebert, S. Andre, J. Jimenez-Barbero and H. Rudiger, *ChemBioChem*, 2004, **5**, 740–764.
47 K. T. Pilobello and L. K. Mahal, *Curr. Opin. Chem. Biol.*, 2007, **11**, 300–305.
48 T. Feizi and W. Chai, *Nat. Rev. Mol. Cell Biol.*, 2004, **5**, 582–588.
49 H. J. Gabius, S. Andre, H. Kaltner and H. C. Siebert, *Biochim. Biophys. Acta*, 2002, **1572**, 165–177.
50 R. A. Laine, *Glycobiology*, 1994, **4**, 759–767.
51 D. B. Werz, R. Ranzinger, S. Herget, A. Adibekian, C. W. von der Lieth and P. H. Seeberger, *ACS Chem. Biol.*, 2007, **2**, 685–691.
52 K. F. Aoki, A. Yamaguchi, Y. Okuno, T. Akutsu, N. Ueda, M. Kanehisa and H. Mamitsuka, *Genome Inform.*, 2003, **14**, 134–143.
53 K. Drickamer and M. E. Taylor, *GenomeBiology*, 2002, **3**, 1034.
54 S. M. Haslam, S. Julien, J. M. Burchell, C. R. Monk, A. Ceroni, O. A. Garden and A. Dell, *Immunol. Cell Biol.*, 2008, **86**, 564–573.
55 S. Doubet and P. Albersheim, *Glycobiology*, 1992, **2**, 505.
56 C. W. von der Lieth, A. Bohne-Lang, K. K. Lohmann and M. Frank, *Briefings Bioinf.*, 2004, **5**, 164–178.
57 N. Taniguchi, A. Ekuni, J. H. Ko, E. Miyoshi, Y. Ikeda, Y. Ihara, A. Nishikawa, K. Honke and M. Takahashi, *Proteomics*, 2001, **1**, 239–247.
58 E. M. Comelli, S. R. Head, T. Gilmartin, T. Whisenant, S. M. Haslam, S. J. North, N. K. Wong, T. Kudo, H. Narimatsu, J. D. Esko, K. Drickamer, A. Dell and J. C. Paulson, *Glycobiology*, 2006, **16**, 117–131.
59 A. V. Nairn, W. S. York, K. Harris, E. M. Hall, J. M. Pierce and K. W. Moremen, *J. Biol. Chem.*, 2008, **283**, 17298–17313.
60 H. Narimatsu, *Glycoconjugate J.*, 2004, **21**, 17–24.
61 H. Ito, A. Kameyama, T. Sato and H. Narimatsu, *Methods Mol. Biol.*, 2009, **534**, 1–10.
62 L. Wells and G. W. Hart, *FEBS Lett.*, 2003, **546**, 154–158.
63 R. Apweiler, H. Hermjakob and N. Sharon, *Biochim. Biophys. Acta*, 1999, **1473**, 4–8.
64 R. G. Spiro, *Glycobiology*, 2002, **12**, 43R–56R.
65 H. Lis and N. Sharon, *Eur. J. Biochem.*, 1993, **218**, 1–27.
66 S. Y. Yu, K. H. Khoo, Z. Yang, A. Herp and A. M. Wu, *Glycoconjugate J.*, 2008, **25**, 199–212.
67 R. Kleene, H. Yang, M. Kutsche and M. Schachner, *J. Biol. Chem.*, 2001, **276**, 21656–21663.
68 C. Bleckmann, H. Geyer, V. Reinhold, A. Lieberoth, M. Schachner, R. Kleene and R. Geyer, *J. Proteome Res.*, 2009, **8**, 567–582.
69 M. L. Reitman and S. Kornfeld, *J. Biol. Chem.*, 1981, **256**, 11977–11980.
70 L. Lang, M. Reitman, J. Tang, R. M. Roberts and S. Kornfeld, *J. Biol. Chem.*, 1984, **259**, 14663–14671.
71 M. Kudo and W. M. Canfield, *J. Biol. Chem.*, 2006, **281**, 11761–11768.
72 P. L. Smith and J. U. Baenziger, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 329–333.
73 P. L. Smith and J. U. Baenziger, *Science*, 1988, **242**, 930–933.
74 T. Satomaa, O. Renkonen, J. Helin, J. Kirveskari, A. Mäkitie and R. Renkonen, *Blood*, 2002, **99**, 2609–2611.
75 G. H. Mir, J. Helin, K.-P. Skarp, R. D. Cummings, A. Mäkitie, R. Renkonen and A. Leppänen, *Blood*, 2009, in press.
76 P. P. Wilkins, R. P. McEver and R. D. Cummings, *J. Biol. Chem.*, 1996, **271**, 18732–18742.
77 Z. S. Kawar, T. K. Johnson, S. Natunen, J. B. Lowe and R. D. Cummings, *Glycobiology*, 2008, **18**, 441–446.
78 P. A. Aeed, J. G. Geng, D. Asa, L. Raycroft, L. Ma and A. P. Elhammer, *Cell Res.*, 2001, **11**, 28–36.
79 M. R. Paquet, S. Narasimhan, H. Schachter and M. A. Moscarello, *J. Biol. Chem.*, 1984, **259**, 4716–4721.
80 A. A. Grey, S. Narasimhan, J. R. Brisson, H. Schachter and J. P. Carver, *Can. J. Biochem.*, 1982, **60**, 1123–1131.
81 R. A. Laine and J. S. Rush, *Adv. Exp. Med. Biol.*, 1988, **228**, 331–347.
82 T. Irimura, T. Tsuji, S. Tagami, K. Yamamoto and T. Osawa, *Biochemistry*, 1981, **20**, 560–566.
83 T. Tsuji, T. Irimura and T. Osawa, *J. Biol. Chem.*, 1981, **256**, 10497–10502.
84 M. Fukuda, A. Dell, J. E. Oates and M. N. Fukuda, *J. Biol. Chem.*, 1984, **259**, 8260–8273.
85 M. Fukuda, A. Dell and M. N. Fukuda, *J. Biol. Chem.*, 1984, **259**, 4782–4791.
86 S. W. Homans, R. A. Dwek, D. L. Fernandes and T. W. Rademacher, *FEBS Lett.*, 1982, **150**, 503–506.
87 K. Yamashita, T. Ohkura, H. Ideo, K. Ohno and M. Kanai, *J. Biochem.*, 1993, **114**, 766–769.
88 K. Yamashita, H. Ideo, T. Ohkura, K. Fukushima, I. Yuasa, K. Ohno and K. Takeshita, *J. Biol. Chem.*, 1993, **268**, 5783–5789.
89 A. H. Plaas, P. J. Neame, C. M. Nivens and L. Reiss, *J. Biol. Chem.*, 1990, **265**, 20634–20640.
90 A. Dell, H. R. Morris, R. L. Easton, M. Panico, M. Patankar, S. Oehniger, R. Koistinen, H. Koistinen, M. Seppala and G. F. Clark, *J. Biol. Chem.*, 1995, **270**, 24116–24126.
91 J. Cisar, E. A. Kabat, M. M. Dorner and J. Liao, *J. Exp. Med.*, 1975, **142**, 435–459.
92 D. Wang, J. M. Hubbard and E. A. Kabat, *J. Biol. Chem.*, 1993, **268**, 20584–20589.
93 D. M. Chipman, J. J. Pollock and N. Sharon, *J. Biol. Chem.*, 1968, **243**, 487–496.
94 S. J. Smith-Gill, J. A. Rupley, M. R. Pincus, R. P. Carty and H. A. Scheraga, *Biochemistry*, 1984, **23**, 993–997.
95 M. R. Pincus, S. S. Zimmerman and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 2629–2633.
96 Y. T. Li, H. Nakagawa, M. Kitamikado and S. C. Li, *Methods Enzymol.*, 1982, **83**, 610–619.
97 H. Ashida, K. Anderson, J. Nakayama, K. Maskos, C. W. Chou, R. B. Cole, S. C. Li and Y. T. Li, *J. Biol. Chem.*, 2001, **276**, 28226–28232.
98 K. M. Anderson, H. Ashida, K. Maskos, A. Dell, S. C. Li and Y. T. Li, *J. Biol. Chem.*, 2005, **280**, 7720–7728.
99 B. C. Zhu and R. A. Laine, *Eur. J. Biochem.*, 1990, **188**, 67–71.
100 R. K. Merkle and R. D. Cummings, *J. Biol. Chem.*, 1987, **262**, 8179–8189.
101 M. S. Nachbar and J. D. Oppenheim, *Methods Enzymol.*, 1982, **83**, 363–368.
102 R. W. Paul, A. H. Choi and P. W. Lee, *Virology*, 1989, **172**, 382–385.
103 P. Luo, G. Canziani, G. Cunto-Amesty and T. Kieber-Emmons, *J. Biol. Chem.*, 2000, **275**, 16146–16154.
104 G. Kretzschmar, A. Toepfer, C. Hüls and M. Krause, *Tetrahedron*, 1997, **53**, 2485–2494.
105 D. Brassart, A. Woltz, M. Golliard and J. R. Neeser, *Infect. Immun.*, 1991, **59**, 1605–1613.
106 C. Mitsuoka, M. Sawada-Kasugai, K. Ando-Furui, M. Izawa, H. Nakanishi, S. Nakamura, H. Ishida, M. Kiso and R. Kannagi, *J. Biol. Chem.*, 1998, **273**, 11225–11233.
107 C. F. Brewer and L. Bhattacharyya, *J. Biol. Chem.*, 1986, **261**, 7306–7310.
108 J. U. Baenziger and D. Fiete, *J. Biol. Chem.*, 1979, **254**, 2400–2407.
109 J. H. Naismith and R. A. Field, *J. Biol. Chem.*, 1996, **271**, 972–976.
110 A. Varki and S. Kornfeld, *J. Biol. Chem.*, 1983, **258**, 2808–2818.
111 H. D. Fischer, K. E. Creek and W. S. Sly, *J. Biol. Chem.*, 1982, **257**, 9938–9943.
112 C. P. Sparrow, H. Leffler and S. H. Barondes, *J. Biol. Chem.*, 1987, **262**, 7383–7390.
113 H. Leffler and S. H. Barondes, *J. Biol. Chem.*, 1986, **261**, 10119–10126.
114 J. Seetharaman, A. Kanigsberg, R. Slaaby, H. Leffler, S. H. Barondes and J. M. Rini, *J. Biol. Chem.*, 1998, **273**, 13047–13052.
115 L. A. Lasky, Sialomucin ligands for selectins: a new family of cell adhesion molecules, in *Proceedings of the 24th International Symposium of the Princess Takamatsu Cancer Research Fund*, Princeton Scientific Publishing Co., Princeton, NJ, 1994, **24**, pp. 81–90.

116 T. Feizi, *Adv. Exp. Med. Biol.*, 2001, **491**, 65–78.
117 M. E. Taylor and K. Drickamer, *Curr. Opin. Cell Biol.*, 2007, **19**, 572–577.
118 A. Leppänen, T. Yago, V. I. Otto, R. P. McEver and R. D. Cummings, *J. Biol. Chem.*, 2003, **278**, 26391–26400.
119 A. Leppänen, S. P. White, J. Helin, R. P. McEver and R. D. Cummings, *J. Biol. Chem.*, 2000, **275**, 39569–39578.
120 A. Leppänen, L. Penttila, O. Renkonen, R. P. McEver and R. D. Cummings, *J. Biol. Chem.*, 2002, **277**, 39749–39759.
121 A. Leppänen, P. Mehta, Y. B. Ouyang, T. Ju, J. Helin, K. L. Moore, I. van Die, W. M. Canfield, R. P. McEver and R. D. Cummings, *J. Biol. Chem.*, 1999, **274**, 24838–24848.
122 S. D. Rosen, *Annu. Rev. Immunol.*, 2004, **22**, 129–156.
123 R. P. McEver, *Glycoconjugate J.*, 1997, **14**, 585–591.
124 B. S. Bochner, R. A. Alvarez, P. Mehta, N. V. Bovin, O. Blixt, J. R. White and R. L. Schnaar, *J. Biol. Chem.*, 2005, **280**, 4307–4312.
125 A. Chandrasekaran, A. Srinivasan, R. Raman, K. Viswanathan, S. Raguram, T. M. Tumpey, V. Sasisekharan and R. Sasisekharan, *Nat. Biotechnol.*, 2008, **26**, 107–113.
126 J. Stevens, O. Blixt, L. M. Chen, R. O. Donis, J. C. Paulson and I. A. Wilson, *J. Mol. Biol.*, 2008, **381**, 1382–1394.
127 S. Gulati, D. F. Smith and G. M. Air, *Virol. J.*, 2009, **6**, 22.
128 Y. Wang, J. Tan, M. Sutton-Smith, D. Ditto, M. Panico, R. M. Campbell, N. M. Varki, J. M. Long, J. Jaeken, S. R. Levinson, A. Wynshaw-Boris, H. R. Morris, D. Le, A. Dell, H. Schachter and J. D. Marth, *Glycobiology*, 2001, **11**, 1051–1070.
129 Y. Takegawa, K. Deguchi, H. Nakagawa and S. Nishimura, *Anal. Chem.*, 2005, **77**, 6062–6068.
130 T. Angata and A. Varki, *Chem. Rev.*, 2002, **102**, 439–469.
131 K. J. Colley, *Neurochem. Res.*, 2008, DOI: 10.1007/s11064-008-9652-6.
132 M. Muhlenhoff, M. Eckhardt and R. Gerardy-Schahn, *Curr. Opin. Struct. Biol.*, 1998, **8**, 558–564.
133 S. V. Evans, B. W. Sigurskjold, H. J. Jennings, J. R. Brisson, R. To, W. C. Tse, E. Altman, M. Frosch, C. Weisgerber and H. D. Kratzin *et al.*, *Biochemistry*, 1995, **34**, 6737–6744.
134 R. Kornfeld and C. Ferris, *J. Biol. Chem.*, 1975, **250**, 2614–2619.
135 R. D. Cummings and S. Kornfeld, *J. Biol. Chem.*, 1982, **257**, 11230–11234.
136 S. Hammarström, M. L. Hammarström, G. Sundblad, J. Arnarp and J. Lönngren, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 1611–1615.
137 T. Feizi, *Immunol. Commun.*, 1981, **10**, 127–156.
138 K. W. Moremen and M. Molinari, *Curr. Opin. Struct. Biol.*, 2006, **16**, 592–599.
139 R. G. Spiro, *J. Biol. Chem.*, 2000, **275**, 35657–35660.
140 D. J. Moloney, A. I. Lin and R. S. Haltiwanger, *J. Biol. Chem.*, 1997, **272**, 19046–19050.
141 D. J. Moloney, L. H. Shair, F. M. Lu, J. Xia, R. Locke, K. L. Matta and R. S. Haltiwanger, *J. Biol. Chem.*, 2000, **275**, 9604–9611.
142 R. G. Spiro, *J. Biol. Chem.*, 1969, **244**, 602–612.
143 B. Schegg, A. J. Hulsmeier, C. Rutschmann, C. Maag and T. Hennet, *Mol. Cell. Biol.*, 2009, **29**, 943–952.
144 K. J. Colley and J. U. Baenziger, *J. Biol. Chem.*, 1987, **262**, 10290–10295.
145 S. Yamada, Y. Okada, M. Ueno, S. Iwata, S. S. Deepa, S. Nishimura, M. Fujita, I. Van Die, Y. Hirabayashi and K. Sugahara, *J. Biol. Chem.*, 2002, **277**, 31877–31886.
146 Y. Tone, L. C. Pedersen, T. Yamamoto, T. Izumikawa, H. Kitagawa, J. Nishihara, J. Tamura, M. Negishi and K. Sugahara, *J. Biol. Chem.*, 2008, **283**, 16801–16807.
147 Y. Guerardel, L. Balanzino, E. Maes, Y. Leroy, B. Coddeville, R. Oriol and G. Strecker, *Biochem. J.*, 2001, **357**, 167–182.
148 C. Robbe-Masselot, A. Herrmann, E. Maes, I. Carlstedt, J. C. Michalski and C. Capon, *J. Proteome Res.*, 2009, **8**, 702–711.
149 K. O. Lloyd and L. J. Old, *Cancer Res.*, 1989, **49**, 3445–3451.
150 R. Kannagi and S. Hakomori, *Adv. Exp. Med. Biol.*, 2001, **491**, 587–630.
151 M. G. Paulick and C. R. Bertozzi, *Biochemistry*, 2008, **47**, 6991–7000.
152 P. Orlean and A. K. Menon, *J. Lipid Res.*, 2007, **48**, 993–1011.
153 M. A. Ferguson, *J. Cell Sci.*, 1999, **112**(pt. 17), 2799–2809.
154 C. A. Bate and D. Kwiatkowski, *Infect. Immun.*, 1994, **62**, 5261–5266.
155 L. Schofield, M. C. Hewitt, K. Evans, M. A. Siomos and P. H. Seeberger, *Nature*, 2002, **418**, 785–789.
156 F. Kamena, M. Tamborrini, X. Liu, Y. U. Kwon, F. Thompson, G. Pluschke and P. H. Seeberger, *Nat. Chem. Biol.*, 2008, **4**, 238–240.
157 N. S. Gandhi and R. L. Mancera, *Chem. Biol. Drug Des.*, 2008, **72**, 455–482.
158 R. Sasisekharan, R. Raman and V. Prabhakar, *Annu. Rev. Biomed. Eng.*, 2006, **8**, 181–231.
159 J. D. Esko and U. Lindahl, *J. Clin. Invest.*, 2001, **108**, 169–173.
160 C. I. Gama, S. E. Tully, N. Sotogaku, P. M. Clark, M. Rawat, N. Vaidehi, W. A. Goddard 3rd, A. Nishi and L. C. Hsieh-Wilson, *Nat. Chem. Biol.*, 2006, **2**, 467–473.
161 J. L. de Paz, C. Noti and P. H. Seeberger, *J. Am. Chem. Soc.*, 2006, **128**, 2766–2767.
162 J. L. de Paz, D. Spillmann and P. H. Seeberger, *Chem. Commun.*, 2006, 3116–3118.
163 C. Noti and P. H. Seeberger, *Chem. Biol.*, 2005, **12**, 731–756.
164 H. A. Orgueira, A. Bartolozzi, P. Schell, R. E. Litjens, E. R. Palmacci and P. H. Seeberger, *Chemistry*, 2003, **9**, 140–169.
165 M. S. Timmer, B. L. Stocker and P. H. Seeberger, *Curr. Opin. Chem. Biol.*, 2007, **11**, 59–65.
166 Z. Shriver, S. Raguram and R. Sasisekharan, *Nat. Rev. Drug Discovery*, 2004, **3**, 863–873.
167 B. Mulloy and R. J. Linhardt, *Curr. Opin. Struct. Biol.*, 2001, **11**, 623–628.
168 U. Lindahl, *Thromb. Haemostasis*, 2007, **98**, 109–115.
169 K. Suzuki, K. Yamamoto, Y. Kariya, H. Maeda, T. Ishimaru, S. Miyaura, M. Fujii, A. Yusa, E. J. Joo, K. Kimata, R. Kannagi, Y. S. Kim and M. Kyogashima, *Glycoconjugate J.*, 2008, **25**, 703–712.
170 J. C. Paulson, O. Blixt and B. E. Collins, *Nat. Chem. Biol.*, 2006, **2**, 238–248.
171 C. F. Brewer, M. C. Miceli and L. G. Baum, *Curr. Opin. Struct. Biol.*, 2002, **12**, 616–623.
172 T. K. Dam, T. A. Gerken, B. S. Cavada, K. S. Nascimento, T. R. Moura and C. F. Brewer, *J. Biol. Chem.*, 2007, **282**, 28256–28263.
173 T. Dam, T. Gerken and C. Brewer, *Biochemistry*, 2009.
174 B. E. Collins, O. Blixt, A. R. DeSieno, N. Bovin, J. D. Marth and J. C. Paulson, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 6104–6109.
175 J. Chatellier, N. Rauffer-Bruyere, M. H. Van Regenmortel, D. Altschuh and E. Weiss, *J. Mol. Recognit.*, 1996, **9**, 39–51.
176 A. J. George, R. R. French and M. J. Glennie, *J. Immunol. Methods*, 1995, **183**, 51–63.
177 A. Azimzadeh and M. H. Van Regenmortel, *J. Immunol. Methods*, 1991, **141**, 199–208.
178 C. J. van Oss, *J. Mol. Recognit.*, 1997, **10**, 203–216.
179 C. A. Chavez, R. N. Bohnsack, M. Kudo, R. R. Gotschall, W. M. Canfield and N. M. Dahms, *Biochemistry*, 2007, **46**, 12604–12617.
180 C. A. King and W. E. Van Heyningen, *J. Infect Dis.*, 1973, **127**, 639–647.
181 J. Holmgren, I. Lonnroth and L. Svennerholm, *Infect. Immun.*, 1973, **8**, 208–214.
182 R. L. Richards, J. Moss, C. R. Alving, P. H. Fishman and R. O. Brady, *Proc. Natl. Acad. Sci. U. S. A.*, 1979, **76**, 1673–1676.
183 W. B. Turnbull, B. L. Precious and S. W. Homans, *J. Am. Chem. Soc.*, 2004, **126**, 1047–1054.
184 A. Schon and E. Freire, *Biochemistry*, 1989, **28**, 5019–5024.
185 P. Cuatrecasas, *Biochemistry*, 1973, **12**, 3547–3558.
186 G. M. Kuziemko, M. Stroh and R. C. Stevens, *Biochemistry*, 1996, **35**, 6375–6384.
187 Y. D. Lobsanov, M. A. Gitt, H. Leffler, S. H. Barondes and J. M. Rini, *J. Biol. Chem.*, 1993, **268**, 27034–27038.
188 D. I. Liao, G. Kapadia, H. Ahmed, G. R. Vasta and O. Herzberg, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 1428–1432.
189 S. H. Barondes, V. Castronovo, D. N. Cooper, R. D. Cummings, K. Drickamer, T. Feizi, M. A. Gitt, J. Hirabayashi, C. Hughes and K. Kasai *et al.*, *Cell*, 1994, **76**, 597–598.
190 J. Hirabayashi, T. Hashidate, Y. Arata, N. Nishi, T. Nakamura, M. Hirashima, T. Urashima, T. Oka, M. Futai, W. E. Muller,

F. Yagi and K. Kasai, *Biochim. Biophys. Acta*, 2002, **1572**, 232–254.
191 M. Sato, N. Nishi, H. Shoji, M. Seki, T. Hashidate, J. Hirabayashi, K. Kasai Ki, Y. Hata, S. Suzuki, M. Hirashima and T. Nakamura, *Glycobiology*, 2002, **12**, 191–197.
192 S. R. Stowell, C. M. Arthur, P. Mehta, K. A. Slanina, O. Blixt, H. Leffler, D. F. Smith and R. D. Cummings, *J. Biol. Chem.*, 2008, **283**, 10109–10123.
193 C. F. Brewer, *Glycoconjugate J.*, 2004, **19**, 459–465.
194 X. Song, B. Xia, S. R. Stowell, Y. Lasanajak, D. F. Smith and R. D. Cummings, *Chem. Biol.*, 2009, **16**, 36–47.
195 S. K. Patnaik, B. Potvin, S. Carlsson, D. Sturm, H. Leffler and P. Stanley, *Glycobiology*, 2006, **16**, 305–317.
196 V. Krishnan, S. M. Bane, P. D. Kawle, K. N. Naresh and R. D. Kalraiya, *Clin. Exp. Metastasis*, 2005, **22**, 11–24.
197 I. Pelletier and S. Sato, *J. Biol. Chem.*, 2002, **277**, 17663–17670.
198 I. Pelletier, T. Hashidate, T. Urashima, N. Nishi, T. Nakamura, M. Futai, Y. Arata, K. Kasai, M. Hirashima, J. Hirabayashi and S. Sato, *J. Biol. Chem.*, 2003, **278**, 22223–22230.
199 T. Takeuchi, K. Hayama, J. Hirabayashi and K. Kasai, *Glycobiology*, 2008, **18**, 882–890.
200 J. Stevens, O. Blixt, J. C. Paulson and I. A. Wilson, *Nat. Rev. Microbiol.*, 2006, **4**, 857–864.
201 O. Blixt, S. Head, T. Mondala, C. Scanlan, M. E. Huflejt, R. Alvarez, M. C. Bryan, F. Fazio, D. Calarese, J. Stevens, N. Razi, D. J. Stevens, J. J. Skehel, I. van Die, D. R. Burton, I. A. Wilson, R. Cummings, N. Bovin, C. H. Wong and J. C. Paulson, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 17033–17038.
202 R. A. Alvarez and O. Blixt, *Methods Enzymol.*, 2006, **415**, 292–310.
203 D. F. Smith and R. D. Cummings, in *Animal Lectins: A Functional View*, ed. G. R. Vasta and H. Ahmed, CRC Press, New York, 2008, pp. 49–62.
204 Z. L. Zhi, A. K. Powell and J. E. Turnbull, *Anal. Chem.*, 2006, **78**, 4786–4793.
205 W. G. Willats, S. E. Rasmussen, T. Kristensen, J. D. Mikkelsen and J. P. Knox, *Proteomics*, 2002, **2**, 1666–1671.
206 H. Tateno, A. Mori, N. Uchiyama, R. Yabe, J. Iwaki, T. Shikanai, T. Angata, H. Narimatsu and J. Hirabayashi, *Glycobiology*, 2008, **18**, 789–798.
207 X. Song, B. Xia, Y. Lasanajak, D. F. Smith and R. D. Cummings, *Glycoconjugate J.*, 2008, **25**, 15–25.
208 S. Fukui, T. Feizi, C. Galustian, A. M. Lawson and W. Chai, *Nat. Biotechnol.*, 2002, **20**, 1011–1017.
209 K. Larsen, M. B. Thygesen, F. Guillaumie, W. G. Willats and K. J. Jensen, *Carbohydr. Res.*, 2006, **341**, 1209–1234.
210 S. Park, M. R. Lee and I. Shin, *Chem. Soc. Rev.*, 2008, **37**, 1579–1591.
211 J. L. de Paz and P. H. Seeberger, *Mol. Biosyst.*, 2008, **4**, 707–711.
212 R. J. Linhardt, J. S. Dordick, P. L. Deangelis and J. Liu, *Semin. Thromb. Hemostasis*, 2007, **33**, 453–465.
213 F. G. Hanisch, C. Hanski and A. Hasegawa, *Cancer Res.*, 1992, **52**, 3138–3144.
214 C. Mitsuoka, N. Kawakami-Kimura, M. Kasugai-Sawada, N. Hiraiwa, K. Toda, H. Ishida, M. Kiso, A. Hasegawa and R. Kannagi, *Biochem. Biophys. Res. Commun.*, 1997, **230**, 546–551.
215 J. C. Yeh, N. Hiraoka, B. Petryniak, J. Nakayama, L. G. Ellies, D. Rabuka, O. Hindsgaul, J. D. Marth, J. B. Lowe and M. Fukuda, *Cell*, 2001, **105**, 957–969.
216 N. Parthasarathy, I. J. Goldberg, P. Sivaram, B. Mulloy, D. M. Flory and W. D. Wagner, *J. Biol. Chem.*, 1994, **269**, 22391–22396.
217 S. Narasimhan, J. C. Freed and H. Schachter, *Carbohydr. Res.*, 1986, **149**, 65–83.
218 J. U. Baenziger and D. Fiete, *J. Biol. Chem.*, 1979, **254**, 9795–9799.
219 B. Walcheck, A. Leppanen, R. D. Cummings, R. N. Knibbs, L. M. Stoolman, S. R. Alexander, P. E. Mattila and R. P. McEver, *Blood*, 2002, **99**, 4063–4069.
220 G. J. Van Dam, A. A. Bergwerff, J. E. Thomas-Oates, J. P. Rotmans, J. P. Kamerling, J. F. Vliegenthart and A. M. Deelder, *Eur. J. Biochem.*, 1994, **225**, 467–482.
221 Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lonngren, J. Arnarp, M. Haraldsson and H. Lonn, *J. Biol. Chem.*, 1983, **258**, 199–202.
222 B. A. Macher, J. Buehler, P. Scudder, W. Knapp and T. Feizi, *J. Biol. Chem.*, 1988, **263**, 10186–10191.
223 D. S. Roseman and J. U. Baenziger, *J. Biol. Chem.*, 2001, **276**, 17052–17057.
224 M. L. Robijn, C. A. Koeleman, M. Wuhrer, L. Royle, R. Geyer, R. A. Dwek, P. M. Rudd, A. M. Deelder and C. H. Hokke, *Mol. Biochem. Parasitol.*, 2007, **151**, 148–161.
225 K. Susuki, N. Yuki and K. Hirata, *J. Neurol. Sci.*, 2001, **185**, 5–9.
226 U. Lindahl, L. Thunberg, G. Backstrom, J. Riesenfeld, K. Nordling and I. Bjork, *J. Biol. Chem.*, 1984, **259**, 12368–12376.
227 U. Lindahl, G. Backstrom, L. Thunberg and I. G. Leder, *Proc. Natl. Acad. Sci. U. S. A.*, 1980, **77**, 6551–6555.
228 J. Choay, M. Petitou, J. C. Lormeau, P. Sinay, B. Casu and G. Gatti, *Biochem. Biophys. Res. Commun.*, 1983, **116**, 492–499.
229 T. Mikami, M. Kashiwagi, K. Tsuchihashi, T. Daino, T. Akino and S. Gasa, *J. Biochem.*, 1998, **123**, 487–491.
230 L. Campos, J. Portoukalian, S. Bonnier, Z. H. Shi, P. Calmard-Oriol, D. Treille and D. Guyotat, *Eur. J. Cancer*, 1992, **28**, 37–41.
231 B. A. Fenderson, A. C. Hahnel and E. M. Eddy, *Dev. Biol.*, 1983, **100**, 318–327.
232 M. Miyake, N. Kohno, E. D. Nudelman and S. Hakomori, *Cancer Res.*, 1989, **49**, 5689–5695.
233 K. Watanabe, S. I. Hakomori, R. A. Childs and T. Feizi, *J. Biol. Chem.*, 1979, **254**, 3221–3228.
234 T. Feizi, R. A. Childs, K. Watanabe and S. I. Hakomori, *J. Exp. Med.*, 1979, **149**, 975–980.
235 J. Le Pendu, R. U. Lemieux and R. Oriol, *Vox Sang.*, 1982, **43**, 188–195.
236 K. Yamashita, K. Totani, Y. Iwaki, M. Kuroki, Y. Matsuoka, T. Endo and A. Kobata, *J. Biol. Chem.*, 1989, **264**, 17873–17881.
237 G. F. Springer and P. R. Desai, *Mol. Immunol.*, 1985, **22**, 1303–1310.
238 J. Dausset, J. Moullec and J. Bernard, *Blood*, 1959, **14**, 1079–1093.
239 G. F. Springer, P. R. Desai and I. Banatwala, *J. Natl. Cancer Inst.*, 1975, **54**, 335–339.
240 E. Klenk and G. Uhlenbruck, *Hoppe- Seyler's Z. Physiol. Chem.*, 1960, **319**, 151–160.
241 B. M. Longenecker, D. J. Willans, G. D. MacLean, S. Selvaraj, M. R. Suresh and A. A. Noujaim, *J. Natl. Cancer Inst.*, 1987, **78**, 489–496.
242 H. Clausen, S. B. Levery, R. Kannagi and S. Hakomori, *J. Biol. Chem.*, 1986, **261**, 1380–1387.
243 M. E. Breimer, H. Karlsson, K. A. Karlsson, K. Nilson, B. E. Samuelsson and N. Stromberg, *Carbohydr. Res.*, 1988, **178**, 111–120.
244 T. Kjeldsen, H. Clausen, S. Hirohashi, T. Ogawa, H. Iijima and S. Hakomori, *Cancer Res.*, 1988, **48**, 2214–2220.
245 A. A. Vyas, O. Blixt, J. C. Paulson and R. L. Schnaar, *J. Biol. Chem.*, 2005, **280**, 16305–16310.
246 U. Galili, B. A. Macher, J. Buehler and S. B. Shohet, *J. Exp. Med.*, 1985, **162**, 573–582.
247 A. K. Nyame, A. M. Leppanen, R. DeBose-Boyd and R. D. Cummings, *Glycobiology*, 1999, **9**, 1029–1035.
248 A. K. Nyame, A. M. Leppanen, B. J. Bogitsh and R. D. Cummings, *Exp. Parasitol.*, 2000, **96**, 202–212.
249 H. C. Gooi, T. Feizi, A. Kapadia, B. B. Knowles, D. Solter and M. J. Evans, *Nature*, 1981, **292**, 156–158.
250 D. Solter and B. B. Knowles, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 5565–5569.
251 K. O. Lloyd, G. Larson, N. Stromberg, J. Thurin and K. A. Karlsson, *Immunogenetics*, 1983, **17**, 537–541.
252 W. W. Young, Jr., H. S. Johnson, Y. Tamura, K. A. Karlsson, G. Larson, J. M. Parker, D. P. Khare, U. Spohr, D. A. Baker, O. Hindsgaul and R. U. Lemieux, *J. Biol. Chem.*, 1983, **258**, 4890–4894.
253 M. Brockhaus, J. L. Magnani, M. Blaszczyk, Z. Steplewski, H. Koprowski, K. A. Karlsson, G. Larson and V. Ginsburg, *J. Biol. Chem.*, 1981, **256**, 13223–13225.
254 K. Fukushima, M. Hirota, P. I. Terasaki, A. Wakisaka, H. Togashi, D. Chia, N. Suyama, Y. Fukushi, E. Nudelman and S. Hakomori, *Cancer Res.*, 1984, **44**, 5279–5285.

255 C. Mitsuoka, K. Ohmori, N. Kimura, A. Kanamori, S. Komba, H. Ishida, M. Kiso and R. Kannagi, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 1597–1602.
256 H. Koprowski, Z. Steplewski, K. Mitchell, M. Herlyn, D. Herlyn and P. Fuhrer, *Somatic Cell Genet.*, 1979, **5**, 957–971.
257 H. Tateno, P. R. Crocker and J. C. Paulson, *Glycobiology*, 2005, **15**, 1125–1135.
258 N. Kimura, K. Ohmori, K. Miyazaki, M. Izawa, Y. Matsuzaki, Y. Yasuda, H. Takematsu, Y. Kozutsumi, A. Moriyama and R. Kannagi, *J. Biol. Chem.*, 2007, **282**, 32200–32207.
259 C. T. Yuen, K. Bezouska, J. O'Brien, M. Stoll, R. Lemoine, A. Lubineau, M. Kiso, A. Hasegawa, N. J. Bockovich and K. C. Nicolaou *et al.*, *J. Biol. Chem.*, 1994, **269**, 1595–1598.
260 Y. Fukushi, E. Nudelman, S. B. Levery, S. Hakomori and H. Rauvala, *J. Biol. Chem.*, 1984, **259**, 10511–10517.
261 W. C. Wang and R. D. Cummings, *J. Biol. Chem.*, 1988, **263**, 4576–4585.
262 R. N. Knibbs, I. J. Goldstein, R. M. Ratcliffe and N. Shibuya, *J. Biol. Chem.*, 1991, **266**, 83–88.
263 N. Shibuya, I. J. Goldstein, W. F. Broekaert, M. Nsimba-Lubaki, B. Peeters and W. J. Peumans, *J. Biol. Chem.*, 1987, **262**, 1596–1601.
264 W. W. Young, Jr., J. Portoukalian and S. Hakomori, *J. Biol. Chem.*, 1981, **256**, 10967–10972.
265 H. Clausen, S. B. Levery, E. Nudelman, S. Tsuchiya and S. Hakomori, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 1199–1203.
266 H. Clausen, S. B. Levery, J. M. McKibbin and S. Hakomori, *Biochemistry*, 1985, **24**, 3578–3586.
267 A. Brown, T. Feizi, H. C. Gooi, M. J. Embleton, J. K. Picard and R. W. Baldwin, *Biosci. Rep.*, 1983, **3**, 163–170.
268 M. J. Seaman, D. G. Chalmers and D. Franks, *Vox Sang.*, 1968, **15**, 25–30.
269 E. I. Park, Y. Mi, C. Unverzagt, H. J. Gabius and J. U. Baenziger, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 17125–17129.
270 C. P. Soh, W. T. Morgan, W. M. Watkins and A. S. Donald, *Biochem. Biophys. Res. Commun.*, 1980, **93**, 1132–1139.
271 D. Blanchard, J. P. Cartron, B. Fournet, J. Montreuil, H. van Halbeek and J. F. Vliegenthart, *J. Biol. Chem.*, 1983, **258**, 7691–7695.
272 J. Finne, M. Leinonen and P. H. Makela, *Lancet*, 1983, **2**, 355–357.
273 J. Finne, U. Finne, H. Deagostini-Bazin and C. Goridis, *Biochem. Biophys. Res. Commun.*, 1983, **112**, 482–487.
274 A. Chiba, K. Matsumura, H. Yamada, T. Inazu, T. Shimizu, S. Kusunoki, I. Kanazawa, A. Kobata and T. Endo, *J. Biol. Chem.*, 1997, **272**, 2156–2162.
275 W. Chai, C. T. Yuen, H. Kogelberg, R. A. Carruthers, R. U. Margolis, T. Feizi and A. M. Lawson, *Eur. J. Biochem.*, 1999, **263**, 879–888.
276 T. Willer, M. C. Valero, W. Tanner, J. Cruces and S. Strahl, *Curr. Opin. Struct. Biol.*, 2003, **13**, 621–630.
277 C. T. Yuen, W. Chai, R. W. Loveless, A. M. Lawson, R. U. Margolis and T. Feizi, *J. Biol. Chem.*, 1997, **272**, 8924–8931.
278 Y. Kuroki, S. Gasa, Y. Ogasawara, M. Shiratori, A. Makita and T. Akino, *Biochem. Biophys. Res. Commun.*, 1992, **187**, 963–969.
279 R. P. Lisak, O. Abramsky, S. H. Dorfman, J. George, M. C. Manning, D. E. Pleasure, T. Saida and D. H. Silberberg, *J. Neurol. Sci.*, 1979, **40**, 65–73.
280 P. Dupouey, A. Billecocq and M. Lefroit, *Immunochemistry*, 1976, **13**, 289–294.
281 J. G. Alvarez, B. T. Storey, M. L. Hemling and R. L. Grob, *J. Lipid Res.*, 1990, **31**, 1073–1081.
282 S. Hakomori, *J. Immunol.*, 1974, **112**, 424–426.
283 F. W. Symington, I. D. Bernstein and S. Hakomori, *J. Biol. Chem.*, 1984, **259**, 6008–6012.
284 K. Nagai, K. Tadano-Aritomi, Y. Niimura and I. Ishizuka, *Glycoconjugate J.*, 2008, **25**, 723–726.
285 H. Leffler, G. C. Hansson and N. Stromberg, *J. Biol. Chem.*, 1986, **261**, 1440–1444.
286 F. W. Symington, B. A. Fenderson and S. Hakomori, *Mol. Immunol.*, 1984, **21**, 877–882.
287 J. T. Clarke and J. A. Embil, *Biochim. Biophys. Acta*, 1979, **582**, 283–294.
288 M. N. Hamers, W. E. Donker-Koopman, M. J. Coulon-Morelec, P. Dupouey and J. M. Tager, *Immunochemistry*, 1978, **15**, 353–358.
289 H. T. Cory, A. D. Yates, A. S. Donald, W. M. Watkins and W. T. Morgan, *Biochem. Biophys. Res. Commun.*, 1974, **61**, 1289–1296.
290 J. Thurin, T. Brodin, B. Bechtel, P. A. Jovall, H. Karlsson, N. Stromberg, S. Teneberg, H. O. Sjogren and K. A. Karlsson, *Biochim. Biophys. Acta*, 1989, **1002**, 267–272.
291 K. Ohyama, Y. Hirabayashi, K. Mizoguchi, M. Nagase, N. Honda and K. Nishimura, *Jpn. J. Exp. Med.*, 1988, **58**, 15–19.
292 N. T. Brodin, J. Thurin, K. A. Karlsson, S. Martensson and H. O. Sjogren, *Int. J. Cancer*, 1989, **43**, 317–326.
293 K. Nagai, D. D. Roberts, T. Toida, H. Matsumoto, Y. Kushi, S. Handa and I. Ishizuka, *J. Biol. Chem.*, 1989, **264**, 16229–16237.
294 K. Nagai, D. D. Roberts, T. Toida, H. Matsumoto, Y. Kushi, S. Handa and I. Ishizuka, *J. Biochem.*, 1989, **106**, 878–886.
295 K. Tadano-Aritomi, T. Kasama, S. Handa and I. Ishizuka, *Eur. J. Biochem.*, 1992, **209**, 305–313.
296 K. Shigeta, Y. Ito, T. Ogawa, Y. Kirihata, S. Hakomori and R. Kannagi, *J. Biol. Chem.*, 1987, **262**, 1358–1362.
297 B. A. Fenderson, P. W. Andrews, E. Nudelman, H. Clausen and S. Hakomori, *Dev. Biol.*, 1987, **122**, 21–34.
298 G. A. Schwarting and D. M. Marcus, *J. Immunol.*, 1977, **118**, 1415–1419.
299 R. Kannagi, S. B. Levery, F. Ishigami, S. Hakomori, L. H. Shevinsky, B. B. Knowles and D. Solter, *J. Biol. Chem.*, 1983, **258**, 8934–8942.
300 L. H. Shevinsky, B. B. Knowles, I. Damjanov and D. Solter, *Cell*, 1982, **30**, 697–705.
301 K. Watanabe and S. Hakomori, *Biochemistry*, 1979, **18**, 5502–5504.
302 A. Ito, S. Saito, T. Masuko, M. Oh-Eda, T. Matsuura, M. Satoh, F. M. Nejad, T. Enomoto, S. Orikasa and S. I. Hakomori, *Glycoconjugate J.*, 2001, **18**, 475–485.
303 M. M. Rapport, H. Schneider and L. Graf, *Biochim. Biophys. Acta*, 1967, **137**, 409–411.
304 R. Laine, C. C. Sweeley, Y. T. Li, A. Kisic and M. M. Rapport, *J. Lipid Res.*, 1972, **13**, 519–524.
305 K. R. Willison, R. A. Karol, A. Suzuki, S. K. Kundu and D. M. Marcus, *J. Immunol.*, 1982, **129**, 603–609.
306 S. Ando, K. Kon, Y. Nagai and T. Yamakawa, *Adv. Exp. Med. Biol.*, 1982, **152**, 71–81.
307 H. Clausen, K. Watanabe, R. Kannagi, S. B. Levery, E. Nudelman, Y. Arao-Tomono and S. Hakomori, *Biochem. Biophys. Res. Commun.*, 1984, **124**, 523–529.
308 H. Clausen and S. Hakomori, *Vox Sang.*, 1989, **56**, 1–20.
309 R. I. Jacobson, N. Kasai, F. F. Richards and R. K. Yu, *J. Neuroimmunol.*, 1982, **3**, 225–235.
310 Y. Hirabayashi, A. Hamaoka, M. Matsumoto, T. Matsubara, M. Tagawa, S. Wakabayashi and M. Taniguchi, *J. Biol. Chem.*, 1985, **260**, 13328–13333.
311 S. K. Kundu, D. M. Marcus and R. W. Veh, *J. Neurochem.*, 1980, **34**, 184–188.
312 C. S. Pukel, K. O. Lloyd, L. R. Travassos, W. G. Dippold, H. F. Oettgen and L. J. Old, *J. Exp. Med.*, 1982, **155**, 1133–1147.
313 E. Nudelman, S. Hakomori, R. Kannagi, S. Levery, M. Y. Yeh, K. E. Hellstrom and I. Hellstrom, *J. Biol. Chem.*, 1982, **257**, 12752–12756.
314 C. Dubois, J. L. Magnani, G. B. Grunwald, S. L. Spitalnik, G. D. Trisler, M. Nirenberg and V. Ginsburg, *J. Biol. Chem.*, 1986, **261**, 3826–3830.
315 T. Ariga, K. Suetake, M. Nakane, M. Kubota, S. Usuki, I. Kawashima and R. K. Yu, *Neurosignals*, 2008, **16**, 226–234.
316 L. D. Cahan, R. F. Irie, R. Singh, A. Cassidenti and J. C. Paulson, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 7629–7633.
317 M. Katano and R. F. Irie, *Immunol. Lett.*, 1984, **8**, 169–174.
318 B. Schwerer, H. Lassmann and H. Bernheimer, *Neuropathol. Appl. Neurobiol.*, 1982, **8**, 217–226.
319 T. Dohi, S. Ohta, N. Hanai, K. Yamaguchi and M. Oshima, *J. Biol. Chem.*, 1990, **265**, 7880–7885.
320 A. Suzuki and T. Yamakawa, *J. Biochem.*, 1981, **90**, 1541–1544.
321 Y. Suzuki, M. Matsunaga, Y. Nagao, T. Taki, Y. Hirabayashi and M. Matsumoto, *Vaccine*, 1985, **3**, 201–203.
322 T. Ariga, K. Kobayashi, Y. Kuroda, R. K. Yu, M. Suzuki, H. Kitagawa, F. Inagaki and T. Miyatake, *J. Biol. Chem.*, 1987, **262**, 14146–14153.

323 K. Nakamura, H. Suzuki, Y. Hirabayashi and A. Suzuki, *J. Biol. Chem.*, 1995, **270**, 3876–3881.
324 S. Furuya, F. Irie, T. Hashikawa, K. Nakazawa, A. Kozakai, A. Hasegawa, K. Sudo and Y. Hirabayashi, *J. Biol. Chem.*, 1994, **269**, 32418–32425.
325 E. Bollensen, H. I. Schipper and A. J. Steck, *J. Neurol.*, 1989, **236**, 353–355.
326 R. Alejandro, F. L. Shienvold, S. A. Hajek, M. Pierce, R. Paul and D. H. Mintz, *J. Clin. Invest.*, 1984, **74**, 25–38.
327 P. Fredman, S. Jeansson, E. Lycke and L. Svennerholm, *FEBS Lett.*, 1985, **189**, 23–26.
328 Y. Hirabayashi, M. Hirota, M. Matsumoto, H. Tanaka, K. Obata and S. Ando, *J. Biochem.*, 1988, **104**, 973–979.
329 I. Kawashima, O. Nakamura and T. Tai, *Mol. Immunol.*, 1992, **29**, 625–632.
330 S. Kusunoki, A. Chiba, Y. Hirabayashi, F. Irie, M. Kotani, I. Kawashima, T. Tai and Y. Nagai, *Brain Res.*, 1993, **623**, 83–88.
331 T. Agui, B. Syuto, K. Oguma, H. Iida and S. Kubo, *J. Biochem.*, 1985, **97**, 213–218.
332 K. Nakamura, H. Kojima, M. Suzuki, A. Suzuki and Y. Tamai, *Eur. J. Biochem.*, 2000, **267**, 5198–5208.
333 H. Yamamoto, S. Tsuji and Y. Nagai, *J. Neurochem.*, 1990, **54**, 513–517.
334 N. Kasai and R. K. Yu, *Brain Res.*, 1983, **277**, 155–158.
335 E. D. Nudelman, S. B. Levery, M. R. Stroud, M. E. Salyan, K. Abe and S. Hakomori, *J. Biol. Chem.*, 1988, **263**, 13942–13951.
336 M. Kyogashima, J. Mulshine, R. I. Linnoila, S. Jensen, J. L. Magnani, E. Nudelman, S. Hakomori and V. Ginsburg, *Arch. Biochem. Biophys.*, 1989, **275**, 309–314.
337 E. D. Nudelman, U. Mandel, S. B. Levery, T. Kaizu and S. Hakomori, *J. Biol. Chem.*, 1989, **264**, 18719–18725.
338 T. Kaizu, S. B. Levery, E. Nudelman, R. E. Stenkamp and S. Hakomori, *J. Biol. Chem.*, 1986, **261**, 11254–11258.
339 M. R. Stroud, S. B. Levery, M. E. Salyan, C. E. Roberts and S. Hakomori, *Eur. J.Biochem.*, 1992, **203**, 577–586.
340 M. R. Stroud, S. B. Levery, E. D. Nudelman, M. E. Salyan, J. A. Towell, C. E. Roberts, M. Watanabe and S. Hakomori, *J. Biol. Chem.*, 1991, **266**, 8439–8446.
341 J. L. Magnani, B. Nilsson, M. Brockhaus, D. Zopf, Z. Steplewski, H. Koprowski and V. Ginsburg, *J. Biol. Chem.*, 1982, **257**, 14365–14369.
342 P. A. Prieto and D. F. Smith, *Arch. Biochem. Biophys.*, 1985, **241**, 281–289.

