

# Application of Microarrays for Deciphering the Structure and Function of the Human Glycome\*

David F. Smith‡ and Richard D. Cummings‡

**Glycan structures were defined historically using multiple methods to determine composition, sequence, linkage, and anomericity of component monosaccharides. Such approaches have been replaced by more sensitive MS methods to profile or predict glycan structures, but these methods are limited in their ability to completely define glycan structures. Glycan-binding proteins, including lectins and antibodies, have been found to have exquisite binding specificities that can provide information about glycan structures. Here, we show glycan-binding proteins can be used along with MS to help define glycan linkages and other determinants in unknown glycans printed as shotgun glycan microarrays. *Molecular & Cellular Proteomics* 12: 10.1074/mcp.R112.027110, 902–912, 2013.**

Although exciting developments in genomics and proteomics during the past 25 years have rapidly advanced our understanding of the structure and function of genes and proteins, they have also exposed the persistent gap in our understanding of the structure and function of the human glycome. The complexity of glycan structures, the potentially large size and dynamic nature of the glycome as well as technical difficulties in glycan sequencing, have helped to create such a disparity (1–6) and will not be reviewed here. However, it is noteworthy that the biomedical community has long been aware of the following: all living organisms have a glycocalyx on their cell surfaces (7–9); expression of a significant percentage of the genome (10, 11), including >700 genes involved in glycan-related processes (12), is required for its synthesis; pathogens invade their hosts by attacking this “sugar” barrier (13–17); and by the late 1960s significant alterations in cell surface glycosylation were recognized as major differences in normal and cancer cells (18, 19). Despite such common facts, this scientific area, recently more formally recognized as glycomics (20–22), has historically been somewhat overlooked by the general biomedical research establishment.

Two major obstacles to the progression of glycomics have been the lack of simple and robust methods for determining

glycan structures (23, 24) and the facile means for glycan synthesis (25, 26), in contrast to the automated methods available for proteins and nucleic acids. The greatest recent advances in glycan sequencing have involved mass spectrometry (27, 28), which now in combination with gene expression studies (29) can link such structural information to the information on biosynthetic pathways that may be helpful in providing clues to glycan sequences by predicting glycan compositions (30, 31). Despite the availability of ultra-sensitive MS methods (32), MS data are limited in their ability to define epimers of hexoses and *N*-acetylhexosamines or to determine positional linkages, and they provide no direct information on anomeric configurations of individual monosaccharides.

*Multiple Approaches for Profiling a Glycome*—Over the years, many approaches were developed to provide information about glycan structures. For example, antibodies and lectins, now generally grouped as glycan-binding proteins (GBPs),<sup>1</sup> have been used in indirect approaches to explore features of cell surface glycans. Antibodies to blood group antigens (ABO and Lewis antigens) (33–35) have been used to identify expression of these determinants on glycoproteins and glycolipids in cells and tissues. Lectins and anti-glycan antibodies have been useful in identifying glycan structures associated with specific glycosylation pathways (36–40), as well as to generate mutant cell lines with altered glycosylation identified by their resistance to toxic lectins (41). Specific GBP binding can provide significant information on the presence of unique monosaccharides (sialic acid and fucose), on the identification of epimers of hexoses (Gal, Man, and Glc), and in many cases anomericity, but they cannot provide complete glycan sequence information. In addition, the quality of lectins and antibodies and their precise specificities are often not well

<sup>1</sup> The abbreviations used are: GBP, glycan-binding protein; Neu5Ac, *N*-acetylneuraminic acid; GSL, glycosphingolipids; AEAB, 2-amino-*N*-(2-aminoethyl)benzamide; UPLC, ultra performance liquid chromatography; Consortium for Functional Glycomics; SGM, shotgun glycan microarray; BBG, bovine brain ganglioside; TGL, tagged glycan library; SNA, *S. nigra* agglutinin; ECL, *E. cristagalli* lectin; GSL-II, *G. simplicifolia* lectin II; Le<sup>a</sup>, Lewis a; Le<sup>x</sup>, Lewis x; MAGS, metadata-assisted glycan sequencing; HMG, human milk glycan; HM-SGM, human milk-shotgun glycan microarray; MVM, minute virus of mice; LSTa, Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-3Gal $\beta$ 1-3GlcNAc $\beta$ 1-3Gal $\beta$ 1-4Glc; DSL, Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-3Gal $\beta$ 1-3(Neu5Ac $\alpha$ 2-6)GlcNAc $\beta$ 1-3Gal $\beta$ 1-4Glc; Con A, Concanavalin A.

From the Department of Biochemistry and the Glycomics Center, Emory University School of Medicine, Atlanta, Georgia 30322

Received December 27, 2012, and in revised form, February 14, 2013

Published, MCP Papers in Press, February 14, 2013, DOI 10.1074/mcp.R112.027110

defined, creating problems in interpretation and replication of experiments.

More direct biochemical approaches to defining glycans involve disassembling the glycocalyx to investigate the structures of specific glycoconjugate classes (42, 43). Free glycans generated by chemical or enzymatic means or directly available in body fluids, such as milk or urine, may be directly analyzed by MS methods, but they are often separated by chromatographic techniques, permethylated, and then characterized by MS methods (44). The amphipathic glycosphingolipids (GSL), which are small relative to glycoproteins and proteoglycans, are isolated from lipid extracts of cells and tissues, separated by nonaqueous chromatographic techniques, permethylated, and characterized directly by MS methods (45). Glycans can be released enzymatically from GSL using recombinant endoglycoceramidases and prior to permethylation and MS sequencing (46). For most common GSL that have a double bond in the sphingosine moiety, ozonolysis followed by degradation at neutral pH eliminates the glycan and provides a simple method for obtaining free GSL-derived glycans that can be reduced, reductively aminated with a variety of tags, permethylated, and analyzed by MS methods (47).

Glycoproteins and proteoglycans may be prepared from biological fluids and aqueous or aqueous-detergent extracts of cells or tissues and then separated by a variety of chromatographic techniques. In cases where a glycomic profile was considered informative, the component glycans may be released by enzymatic or chemical means, such as by treatment with proteases followed with endoglycosidases, e.g. peptide-*N*-glycosidase-F and *O*-glycanase, to release *N*-glycans and the Ser/Thr-linked core 1 *O*-glycan, respectively, or treated with base, e.g. NaOH or hydrazine, to release Ser/Thr-linked *O*-glycans by  $\beta$ -elimination using reductive or nonreductive methods (48). The resulting free glycans, like those derived from GSL, are typically either directly derivatized by fluorescent tags, e.g. 2-aminobenzamide (49) or 2-amino-*N*-(2-aminoethyl)benzamide (AEAB) (50) for separation by HPLC (51), or directly reduced, permethylated, and analyzed by MS analysis to obtain structural information (1, 5, 27, 28, 42, 52). Thus, methods to obtain the glycans comprising the glycoconjugates of a glycome are generally available, but they all have specific limitations.

*Glycan Sequencing Approaches Can Be Combined to Aid in Defining the Glycome*—Current ultrasensitive MS methods are routinely used for predicting glycan composition and sequence, and if sufficient standards and sophisticated tandem MS technologies are available to reproducibly identify unique fragment ions, detailed glycan structural analyses are possible by highly skilled investigators. Such complex structural studies are time-consuming and difficult but can ultimately help to provide the list of glycan structures comprising a glycome. Although such a list can provide the “road map” for the synthetic organic chemists to synthesize the glycans com-

prising a glycome, the information alone has intrinsically little value, and the unfortunate lack of facile synthetic methods put the “time until we have a glycome” far into the future. Thus, if glycomics is to evolve along the lines of genomics and proteomics, simple and robust methods that complement MS and permit the “capture” of natural glycans for functional studies are also required. Such methods can evolve from the technologies that were classically performed on pure compounds, which would also provide access to the glycans for functional analysis. Although NMR spectrometry has long been considered a definitive tool for identifying complete glycan structures (53, 54) and is probably absolutely required for microbially derived complex polysaccharides (55), its application is limited because few naturally occurring glycans of interest are available in sufficient quantities for NMR analysis.

A variety of separation techniques have been applied to glycan analyses with great success, and if carefully monitored and validated, these methods could provide significant structural information based on comparison of chromatographic properties with known standards, as well as direct analyses by MS. Examples of significant success in these areas include ultra performance liquid chromatography (56), HPLC-MALDI-TOF MS (57), high performance anion-exchange chromatography (58–60), two-dimensional HPLC (61–63), high performance capillary electrophoresis (42), and fluorophore-assisted carbohydrate electrophoresis (64, 65). All of these methods rely on the high resolution of the separation techniques and the sensitivity of detection of separated glycans; however, all require the availability of defined standard glycans for comparison. Methods employing multiple chromatographic strategies, including hydrophilic interaction liquid chromatography (66), can be useful in comparative analysis of glycomes, as elegantly applied to serum glycoprotein glycosylation (23). However, most of these approaches suffer from the possibility that undefined glycans can co-purify with known structures and thus severely limit the discovery of previously unidentified glycans.

Serial lectin affinity chromatography (37, 67, 68) was an early strategy for analyzing glycans based on knowledge of specific structural features in glycans required for their interactions with lectins. In this approach, glycans either metabolically radiolabeled (69), end radiolabeled, or fluorescently labeled were analyzed by their chromatographic properties and affinity to specific lectins and antibodies. In these studies, glycan structures were determined by their specific binding to defined lectins and antibodies, co-chromatography compared with standards, and the use of highly purified, specific endo- and exo-glycosidase digestions to monitor changes in chromatographic behavior (70). Anomeric configurations and linkage positions of monosaccharides in a glycan were assessed based on susceptibility or resistance to specific enzyme degradation. This approach using HPLC-based methods is clearly useful for analyzing extremely small amounts of material (71). In some cases it has been possible to define

branching of radiolabeled *N*-glycans using acetolysis (72) and linkage analyses on extremely small quantities of glycans by subjecting metabolically radiolabeled glycans to permethylation, hydrolysis, acetylation, and identification of methylated monosaccharides by gas chromatography with detection of radioisotope in the eluted gas (67, 73). Although these approaches are laborious and time-consuming, they can generate information on the detailed structure of glycans at levels of sensitivity that are equivalent, if not superior, to some of the more sophisticated methods used today. An example of the use of such lectins in glycomics studies is the use of the plant lectin LPHA, the leukoagglutinin from *Phaseolus vulgaris* (red kidney bean). This lectin binds to a unique set of isomers of complex-type *N*-glycans containing branched mannose with *N*-acetylglucosamine units linked  $\beta$ 1–2 and  $\beta$ 1–6 to  $\alpha$ -linked mannose (the so-called 2,6-branched mannose) (74). In cells (75) and mice (76) lacking the  $\beta$ 1–6 *N*-acetylglucosaminyltransferase, the binding of LPHA is lost. Thus, this simple plant lectin is exquisitely suited to provide detailed information about *N*-glycan branching and expression and to distinguish *N*-glycan isomers and glycoproteins containing such structures in glycoproteomic analyses (77).

*Defined Glycan Microarrays Provide Clues to Understanding GBP Function*—The publicly available defined glycan microarray developed by the National Institutes of Health-funded Consortium for Functional Glycomics (CFG) has made a major impact on the advancement of functional glycomics analysis. After the analysis of hundreds of GBPs on this array, investigators in many areas are beginning to appreciate access to glycans in a microarray format that can generate information on GBP function. Because it is not possible to amplify a glycome by a method like PCR, immobilized glycans on an array are analogous to amplified products of the genome, *i.e.* oligonucleotides, genes, gene fragments, and recombinant proteins, that make them available for functional studies.

The information derived by interrogating defined glycan microarrays with GBPs has led to important discoveries of GBP function. For example, the observation that galectins-3, -4, and -8 at physiological concentrations bind to human blood group glycans (78) suggested that they may play an innate immune role in humans. Humans are limited in generating the effective adaptive immune responses to blood group antigens recognized as self while being exposed to microorganisms expressing self-like blood group-related glycans. This hypothesis was confirmed by demonstrating that these galectins could not only bind to bacteria expressing blood group-related antigens but that they were in many cases bactericidal (79). Thus, data from defined glycan arrays represent a single piece of information regarding the glycan-binding specificity of a GBP. The GBP specificity is an important property that can be the basis for generating new hypotheses regarding GBP function. Knowledge of other important biological properties of a GBP, such as biological

activity, tissue expression levels, and subcellular location, are normally required to rigorously define the GBP function.

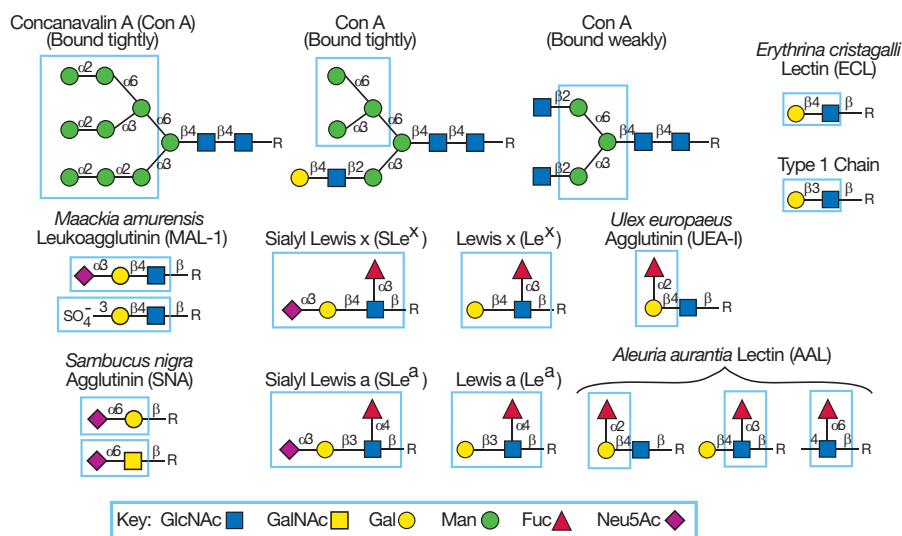
In our approach to defining a glycome, we and others have derivatized free glycans derived from cells and tissues with a bifunctional tag that is fluorescent and also carries a free amino group (2,6-diaminopyridine (80) and more recently AEAB (50) or 2-aminobenzamide (81)). Fluorescence provides a method for detecting glycans during their purification, and the amino function provides a reactive center to immobilize glycans for functional analyses on glycan microarrays or other solid phases.

*Shotgun Glycan Microarrays Define Biologically Relevant Glycans*—Defined glycan microarrays are limited to the glycans we have available for printing on an array, and in many instances a defined array is missing important glycan structure(s) required to define a glycan specificity or epitope. Theoretically, a defined glycan array comprising all member glycans within the human glycome would allow us to define the specificity of any biologically relevant GBP. Currently, the CFG-defined glycan microarray with ~600 glycans represents <10% of the number of glycans estimated to comprise the human glycome. Defining and making the human glycome available in a format that can be screened by GBPs should be an important mission of the field of glycomics. One approach to address this goal will be to determine the detailed structures of the human glycome, which is probably composed of glycans containing at least 10,000 determinants (4), and then have them generated by synthetic chemists so that the human glycome would be available for functional analyses. Based on current technology, however, it is unrealistic to imagine either that the detailed structure of the human glycome can be accomplished in a reasonable time or that such a large number of glycans could be chemically synthesized. It is probably more reasonable to believe that a few thousand glycans can be synthesized by chemical and chemo-enzymatic synthesis during the next decade, but a lot of time may be focused on synthesizing structures of little biological interest. To direct the efforts of such a project, we have proposed shotgun glycomics as a method to identify physiologically or biologically relevant glycans that are screened as potential glycan ligands for GBPs of interest (82).

We developed shotgun glycomics to approach the definition of the function *and* structure of the human glycome in a relatively high throughput manner using naturally expressed glycans. We used nanoscale methods to isolate glycans from natural sources and prepare glycan libraries for direct studies of both their structure and function in terms of GBP recognition. This approach focuses sequencing efforts on functionally relevant glycans recognized by a GBP and results in libraries of naturally occurring glycans that can be archived and retrieved for future studies. In our first example of a shotgun glycan microarray (SGM) (82), ozonolysis of the sphingosine portion of a mixture of GSLs generated free aldehydes that readily reacted by reductive amination to create fluorescent GSL derivatives with a primary amino group. The mixture of

**FIG. 1. Glycan determinants defined by glycan-binding proteins.**

The determinants defined by six commercially available lectins and five monoclonal antibodies were determined by analysis on the defined glycan microarray provided by the Consortium for Functional Glycomics. The determinants, whose presence on an array can be defined by positive signals in an array analysis, are outlined in each structure. *Con A*, concanavalin A.



GSL derivatives composed of bovine brain gangliosides (BBG) was resolved by two-dimensional HPLC into 40 individual derivatives that make up a BBG-tagged glycan library (TGL). The derived glycolipids were quantified based on their fluorescence, characterized by MALDI-TOF/TOF analysis, and printed at equimolar concentrations on *N*-hydroxysuccinimide-derivatized slides. We interrogated the SGM for antibodies to GSL in serum from individuals that had been diagnosed with Lyme disease. One BBG fraction showed significant recognition by sera from individuals with Lyme disease compared with control sera, and the GSL derivative was retrieved from the TGL for further structural characterization. The predicted structure was GD1b-lactone, which is a relatively rare component of brain tissues and melanoma cells (82). Although these data supported additional studies with a larger set of serum samples, they more importantly demonstrated the utility of an SGM for identifying a relevant but rare GSL in a crude fraction of BBG. Similar SGMs from human erythrocytes and a cultured human cell line are described in this report (82), demonstrating that this technology can generate purified natural glycans for defining the human glycome.

In this strategy for human glycomics analysis, the term “shotgun” refers to the fact that glycans are prepared from specific cells or tissues and differs from shotgun genomics in that it does not propose to directly sequence all of the component member glycans in the TGL, but to prioritize structural efforts and identify glycans to be synthesized by chemists for expanding the defined array. As structural definition progresses, the number of defined structures on the SGM will increase, and ultimately the entire TGL may be defined. The shotgun glycomics approach may be applied to any organism, and it might be particularly important for model organisms such as *Caenorhabditis elegans*, *Drosophila*, and zebrafish.

*Metadata-assisted Glycan Sequencing, a Glycomics Approach Based on MS Analysis of TGLs and Defined GBP Binding to SGMs*—The TGL generated from a tissue or orga-

nism represents a significant component of a glycome, and each glycan fraction has an associated mass based on MALDI-TOF analysis carried out prior to printing the SGM. Defined GBPs, e.g. plant and animal lectins and anti-glycan antibodies, provide a rich source of reagents for detecting unique glycan determinants among the glycans printed on an SGM. To validate the printing process, we normally interrogate the SGM with defined GBPs to be sure glycans were printed. Thus, in the process of validating the printing of an SGM with these reagents, we generated significant structural information. Fig. 1 provides a description of the unique glycan determinants of a small selection of commercially available lectins and antibodies used to introduce this approach. Concanavalin A is capable of detecting *N*-glycans due to its specificity for  $\alpha$ -linked Man in branched  $\text{Man}_{\alpha 1-3}(\text{Man}_{\alpha 1-6})\text{Man}_{\alpha 1-R}$  (83, 84), as well as its weaker interaction with the internal trimannosyl core of bi-antennary *N*-glycans, but not with tri- or tetra-antennary or bisected *N*-glycans. *Sambucus nigra* agglutinin (SNA) is generally considered specific for  $\text{Neu5Ac}_{\alpha 2-6}\text{Gal}\beta 1-4\text{GlcNAc}$  (85, 86), but it binds better to  $\text{Neu5Ac}_{\alpha 2-6}\text{Gal}\beta 1-4\text{GlcNAc}\beta 1-3\text{Man}_{\alpha 1-3}\text{Man}$  sequences on *N*-glycans than to the terminal sequence on the six branch of bi-antennary *N*-glycans (87). *Maackia amurensis* lectin I detects  $\text{Neu5Ac}_{\alpha 2-3}\text{Gal}\beta 1-4\text{GlcNAc}$  (88), and *Erythrina cristagalli* lectin (ECL) is specific for  $\text{Gal}\beta 1-4\text{GlcNAc}$  (89). Among the fucose-binding lectins, the examples provided are *Aleuria aurantia* lectin, which has a rather broad specificity for  $\alpha$ -linked Fuc (90), whereas *Ulex europaeus* agglutinin is specific for H-antigen ( $\text{Fuc}_{\alpha 1-2}\text{Gal-R}$ ) (91). A variety of anti-glycan monoclonal antibodies are now commercially available, and the numbers are growing rapidly (4); five examples of different antibody specificities among the dozens defined to date are shown in Fig. 1.

MALDI-TOF data provide the molecular masses and the composition of the glycans printed on the SGM. The binding



Predicted Lectin/Antibody Binding to Glycan Immobilized on Microarray

10 possible isobaric N-glycan structures <sup>1</sup>	Before / After Neuraminidase										
	MAL-1	SNA	ECL	Type 1	Le <sup>x</sup>	SLe <sup>x</sup>	Le <sup>a</sup>	SLe <sup>a</sup>	UEA-I	AAL	ConA
#1	+ / -	- / -	+ / +	- / -	- / -	- / -	- / -	- / -	- / -	+ / +	+ / +
#2	+ / -	- / -	- / +	+ / +	- / -	- / -	- / -	- / -	- / -	+ / +	+ / +
#3	- / -	+ <sup>2</sup> / -	+ / +	- / -	- / -	- / -	- / -	- / -	- / -	+ / +	+ / +
#4	- / -	+ / -	+ / +	- / -	- / -	- / -	- / -	- / -	- / -	+ / +	+ / +
#5	+ / -	- / -	- / +	- / -	- / -	- / -	- / -	- / -	+ / +	+ / +	+ / +
#6	+ / -	- / -	- / +	- / -	+ / +	- / -	- / -	- / -	- / -	+ / +	+ / +
#7	+ / -	- / -	- / +	- / -	- / -	- / -	+ / -	- / -	- / -	+ / +	+ / +
#8	- / -	- / -	+ / +	- / -	- / +	+ / -	- / -	- / -	- / -	+ / +	+ / +
#9	- / -	- / -	+ / +	- / -	- / -	- / -	- / -	+ / -	- / -	+ / +	+ / +
#10	- / -	+ <sup>2</sup> / -	- / +	- / -	- / -	- / -	+ / +	- / -	- / -	+ / +	+ / +

Key: Gal GlcNAc Man Neu5Ac Fuc

<sup>1</sup> Glycan Mass = 2077.745 (monoisotopic) for N-glycan released by N-glycanase (PNGase F). Predicted composition Neu5Ac(1)Hex(5)HexNAc(4)dHex(1).

<sup>2</sup> SNA binding to this isomer #3 is relatively weak for the sequence Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc $\beta$ 1-2Man $\alpha$ 1-6, compared to binding to isomer #4, where Neu5Ac $\alpha$ 2-3 is on the other branched mannose (Man $\alpha$ 1-3).

**FIG. 2. Predicted and unique lectin/antibody binding patterns to hypothetical isobaric N-glycans immobilized on a hypothetical glycan microarray used for structural analysis.** Ten hypothetical N-glycans (1–10) are listed, and the predicted binding patterns for the hypothetical microarray of the lectins and anti-glycan antibodies described in Fig. 1 are shown. The + or – indicate positive or negative binding, and each data point is divided to provide the results of binding with no treatment (*above the slash*) and the results of binding after treatment with nonspecific neuraminidase (*below the slash*).

patterns of the different lectins can provide extensive structural information and can be extremely useful for differentiating glycans that have the same mass but a different arrangement of monosaccharides. In Fig. 2, we summarize the data that would be generated from a glycan microarray of 10 isobaric, bi-antennary N-glycans, whose structures are shown with the pattern of binding of the 11 GBP specificities described in Fig. 1. Interestingly, no two patterns were identical despite the fact that all of the glycans were biantennary N-glycans with the same composition.

For example, glycans 6 and 7 (Fig. 2) differ only by the linkages of Gal and Fuc in the 3-branch of the bi-antennary structure, but they are readily distinguished by their binding with anti-Lewis a (Le<sup>a</sup>) and anti-Lewis x reagents. The positive MAL binding together with antibody binding confirmed the branched structure because these reagents are specific for terminal structures. The data, however, cannot determine on which branch each determinant resides. After neuraminidase digestion, ECL binding was positive indicating that the  $\alpha$ 3-linked sialic acid is located on Gal $\beta$ 1-4GlcNAc. Space does

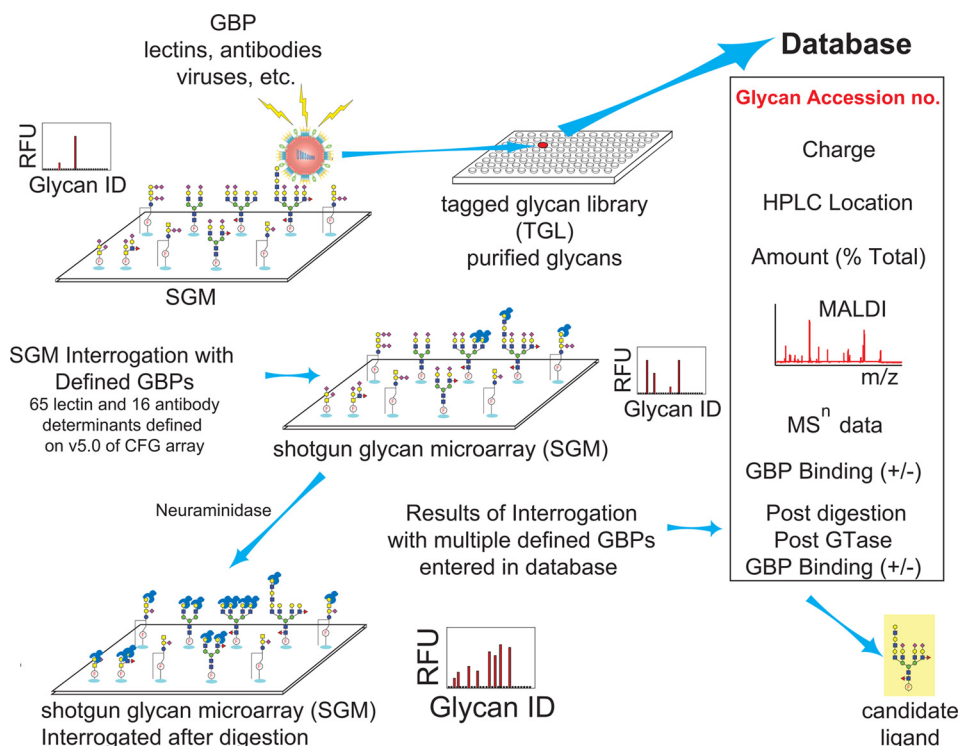


FIG. 3. **Metadata-assisted glycan sequencing is an extension of the shotgun glycan microarray concept.** Beginning with the generation of the TGL, each glycan is assigned an accession number and printed on the array, and the metadata are collected for each glycan and stored in a database. Pre-printing information can include the following: number of negative charges based on ion-exchange chromatography; location of the glycan in the two-dimensional HPLC separation profiles; percentage of total glycomes that each glycan represents; MALDI-TOF data to provide information on purity, composition, additional MS<sup>n</sup> data as obtained; defined GBP binding before and after exoglycosidase digestion; and any other information deemed useful regarding the nature of the glycan.

not permit the interpretation of each data point, but a large amount of data can be associated with each glycan on this example array. These metadata can be compiled in a database and used for predicting monosaccharide sequence and detailed structures of each individual glycan. We have termed this approach metadata-assisted glycan sequencing (MAGS).

MAGS is based on the analyses of many replicate arrays of undefined glycans (SGMs) that are interrogated by many different GBPs (Fig. 3). As the SGM is interrogated with defined GBPs as well as GBPs whose specificity and function are unknown, a database continues to be populated with information on each glycan. When a glycan is determined to be biologically relevant based on a binding event, additional information may be obtained by retrieving the glycan from the TGL for further analysis, although structural information on the entire glycome on the SGM can be addressed by evaluating the binding profile of defined GBPs before and after specific *in situ* exoglycosidase digestion on the arrays (92).

**Analysis of Human Milk SGM by MAGS**—A variety of techniques has been applied to solve the structures of the complex mixture of isomeric glycans found in human milk (93). We elected to take a functional approach to the human milk-soluble glycome using shotgun glycomics (92). The neutral, monosialyl, and disialyl glycans of a milk sample were tagged

with AEAB, separated by two-dimensional HPLC into 127 nearly homogeneous but not fully characterized glycans that made up the human milk TGL. During the production of the TGL, we analyzed each fraction by MALDI-TOF analysis and accumulated data for each glycan. The TGL was printed as a microarray of 127 glycans ( $n = 4$ ) on an *N*-hydroxysuccinimide-derivatized microscope slide to produce the “human milk shotgun glycan microarray” or HM-SGM. To demonstrate the utility of the human milk glycan (HMG) array, we interrogated it with a variety of well characterized GBPs, including lectins and specific anti-glycan antibodies. No significant binding was observed with concanavalin A, *Vicia villosa* lectin, *Griffonia simplicifolia* lectin II (GSL-II), and *M. amurensis* lectin I, consistent with the absence of mannose, terminal GalNAc, terminal GlcNAc, and terminal Neu5Ac $\alpha$ 2-3Gal $\beta$ 1-4GlcNAc, respectively, in human milk (92). The other six lectins, *A. aurantia* lectin, SNA, *Lotus tetragonolobus* lectin, *U. europaeus* agglutinin, *Ricinus communis* agglutinin I, and ECL, exhibited binding to many HMGs on the array.

In interrogations to identify the function of HMGs, we discovered a number of interesting features of these glycans. Some glycans contain epitopes for the monoclonal antibodies TRA-1-60 and TRA-1-81, which are specific for biomarkers of human embryonic pluripotent stem cells (94). Other specif-

**A** Lectin and Antibody Binding to Standards and Unknown Disialyl HMG

glycans analyzed	SNA	ECL	AAL	GSL-II	Le <sup>a</sup>	Type 1	SLe <sup>a</sup> LSTa
unknown <sup>1</sup> 	-	+	-	-	-	-	+
#1 	-	+	-	-	-	-	-
#2 	+	-	-	-	-	-	-
#3 	-	-	-	-	-	-	-
#4 	-	-	-	-	-	+	-
#5 	-	-	-	+	-	-	-
#6 	-	+	-	-	-	-	-
#7 	-	-	-	-	-	-	+
#8 	-	-	-	-	-	-	-
#9 	+	-	-	-	-	-	-
#10 	-	-	-	-	-	-	+

**B** Lectin and Antibody Binding to Standards and Unknown Disialyl HMG before and after exoglycosidase digestions

treatment	no Enzyme	$\alpha$ 2,3 NA	$\alpha$ 2,3+6 NA	$\alpha$ 2,3 NA	$\alpha$ 2,3 NA	$\alpha$ 2,3+6 NA	$\alpha$ 2,3+6 NA
				$\beta$ 1,3 Galase	$\beta$ 1,4 Galase	$\beta$ 1,3 Galase	$\beta$ 1,4 Galase
predicted structures							
lectins/antibodies							
SNA	-	-	-	-	-	-	-
ECL	+	+	+	+	-	+	-
GSL-II	-	-	-	-	-	+	-
anti-Type 1	-	-	+	-	-	-	+
anti-LSTa/ SLe <sup>a</sup>	+	-	-	-	-	-	-

Key: Gal (yellow circle), GlcNAc (blue square), Glc (green circle), Neu5Ac (purple diamond)

<sup>1</sup> Glycan Mass = 1818.174 (monoisotopic). Predicted composition Neu5Ac(2)Hex(4)HexNAc(2).

ically sialylated glycans are bound by fluorescently labeled influenza A virus and minute virus of mice (MVM), suggesting that HMGs may function as receptor decoys in an innate defense mechanism against potential pathogens (92). Overall, influenza A bound to eight glycans; MVM bound to six glycans, and the TRA-1 antibodies bound to six different glycans on the HM-SGM (total of 20 different glycans). Whereas molecular mass data provided compositions of the natural glycan ligands bound by these potential pathogens, more detailed MS<sup>7</sup> data were unable to definitively solve the structures. To obtain more decisive structural characterizations and demonstrate the utility of the MAGS, we retrieved 22 functionally identified glycans from the TGL of the HM-SGM and printed them with 17 defined milk glycan standards. Individual arrays were interrogated with eight lectins and five anti-glycan monoclonal antibodies that had been analyzed on the CFG glycan microarray to confirm their specificity and binding activity. Such an analysis is essential in these studies due to the lack of vendor quality control of these reagents. After obtaining the initial patterns of binding, the subarrays were subjected to digestion with specific exoglycosidases either independently or in combination and subsequently interrogated again with the appropriate lectins or anti-glycan antibodies with positive and negative binding indicating the presence or absence of the corresponding determinants. The results of these studies are summarized in heat maps along with detailed descriptions of the logic behind the assignments of predicted structures of the functionally identified glycans (92). Because of space limitations, the detailed analysis of a single disialylated glycan identified as a glycan ligand bound by MVM is provided here as an example of a structural determination using the MAGS approach. However, the structural data of 22 HMGs on the array were obtained simultaneously.

Several assumptions can be made regarding HMGs based on previous studies (93, 95, 96). They all have lactose as a reducing disaccharide and are composed of a single glucose residue with Gal and GlcNAc present in linear or branched sequences of Gal $\beta$ 1–3/4GlcNAc (LacNAc). The GlcNAc is linked  $\beta$ 1–3 to Gal in linear glycans with branches occurring when GlcNAc is attached  $\beta$ 1–6 to Gal. These core glycans are then substituted with  $\alpha$ -linked Fuc and  $\alpha$ -linked Neu5Ac to make up an extremely complex mixture of isomeric and isobaric glycans. The results of interrogation of the disialylated glycan and 10 standard milk glycans with a selection of de-

defined lectins and antibodies are shown in Fig. 4A, and the results of interrogation after exoglycosidase and sequential exoglycosidase digestions are shown in Fig. 4B. The unknown glycan (predicted structure shown in Fig. 4, A and B) is an octasaccharide (mass = 1818.174) composed of three residues of Gal, one Glc, two residues of GlcNAc, and two residues of Neu5Ac based on the fact that it originated from human milk. ECL and the anti-SLe<sup>a</sup>/LSTa antibody bind the unknown glycan, indicating that it is a branched structure containing a terminal type 2 glycan (ECL-positive) and a terminal LSTa determinant (absence of fucose excludes the possibility of SLe<sup>a</sup>). ECL binding was weak (data not shown) and was presumably due to the steric effect of the sialylated branch, because the ECL binding signal increased by 3-fold after neuraminidase treatment (data not shown), and one branch must be a type 1 structure, which is not bound by ECL. Because digestion with  $\alpha$ 2–3-neuraminidase leads to no change in ECL binding (data not shown), we assume that the other sialic acid must be  $\alpha$ 2–6-linked and continues to block the ECL binding even after removal of the  $\alpha$ 2–3-sialic acid (Fig. 4B). In addition, anti-type 1 antibody binding is observed only after removal of all the sialic acid by nonspecific neuraminidase. Thus, the disialylated glycan is predicted to contain one terminal type 2 and one terminal disialyl LNT (DSL). In addition, no GSL-II binding is observed after  $\beta$ 1–4 galactosidase treatment (Fig. 4B), because GSL-II does not bind GlcNAc $\beta$ 1–6Gal, which would be exposed by sequential  $\alpha$ 2–3-neuraminidase and  $\beta$ 1–3-galactosidase digestion (Fig. 4B). However, sequential nonspecific neuraminidase and  $\beta$ 1–3-galactosidase treatment did lead to strong binding to GSL-II, consistent with the predicted structure. The low SNA binding is consistent with the lack of SNA binding to the DSL standard (Fig. 4A). Taken together, the MAGS analysis permits a relatively conclusive prediction that the isolated glycan ligand bound by MVM is as shown in the unknown structure in Fig. 4A.

Using this conceptual approach, we were able to propose structures, including most of the linkage positions and anomeric configurations for the glycans that bound TRA-1 antibodies, influenza A, and MVM (92). Some redundancy of structures was observed due to the overlap of glycan fractions collected during multidimensional chromatography. Nevertheless, the correlation of proposed structures with the function defined by antibody or virus binding provided information on the glycans in human milk that are related to

**FIG. 4. Example of MAGS of a single human milk glycan selected for structural analysis based on its binding function.** A, lectin and antibody binding to standards and an unknown disialyl human milk glycan. The predicted structure of the unknown glycan identified as a glycan ligand for MVM from human milk is shown at the upper left above a list of 10 glycan standards obtained from human milk. The binding patterns for four defined lectins (SNA, ECL, *A. aurantia* lectin, and GSL-II) and three monoclonal antibodies (anti-Le<sup>a</sup>, anti-type I glycan (Gal $\beta$ 1–3GlcNAc), and anti-Sialyl Le<sup>a</sup>/LSTa) from individual microarray analyses are indicated as either positive (+) or negative (–), indicating the presence or absence of the determinants as defined in Fig. 1. B, lectin and antibody binding to standards and an unknown disialyl human milk glycan before and after exoglycosidase digestion. The predicted structures of the unknown glycan and exoglycosidase products of the enzyme treatments are shown with the patterns of lectin and antibody binding from individual microarray analyses indicated as either + or –.



embryonic stem cell-specific epitopes and potential receptors for viruses.

Specific lectins are commonly used for the immunohistochemical localization of specific glycan structures in cells and tissues, and with MAGS we simply apply this approach to libraries of structurally undefined glycans with the ability to obtain additional material from the reserved TGL entries for deeper structural characterization. Although the MAGS approach to determining glycan structure alone cannot unequivocally identify a glycan, it can certainly complement MS in glycan structural determination, because it incorporates additional chemical and biochemical approaches used to define glycan structure, including digestions with specific exo-glycosidases, and specific lectin and antibody interactions. The concept of collecting metadata on structural analysis arrays is rational, and it is not unreasonable to consider the possibility that this approach with miniaturization, increased detection sensitivity, and automation could be developed as a high throughput approach to glycomics analysis.

Although MAGS represents a novel approach to assist in high throughput sequencing of glycans, it is highly dependent upon the availability of well characterized GBPs, such as antibodies and lectins, of defined specificity. Although many such GBPs are available, the quality control of commercial reagents is sometimes lacking; thus, researchers unaware of this may draw incorrect conclusions without independent verification of the activity and specificity of the commercial GBP. Certainly, a major limitation of the MAGS approach here is the availability and costs of large numbers of quality-controlled GBPs. Another important aspect of this approach is the requirement of high quality purified, specific exoglycosidases that are active and efficient on immobilized glycans on the glycan arrays being used for MAGS. In our studies, we have observed that the concentration of enzyme required to completely hydrolyze a terminal residue from a printed glycan is significantly higher than the amount of enzyme required to accomplish the same result in solution. Finally, the analysis and interpretation of the binding patterns of these defined glycans are also relatively new areas carried out by a relatively small number of laboratories that require specialized equipment for printing and scanning glycan microarrays. These techniques continue to be refined as more and varied presentations of glycans on microarrays are developed.

Bioinformatics will be an absolute requirement for developing and maintaining the databases generated by MAGS and ultimately mining the information to obtain useful data. The simultaneous analyses of hundreds of unknown glycan structures using multiple analyses of replicate printed arrays before and after single or sequential enzyme digests will result in massive amounts of data that can only be handled computationally. Our initial approach to bioinformatics has developed out of our need to process data from the interaction of unknown GBPs with defined glycan arrays with the purpose to automate the discovery of GBP-binding motifs and define the

glycan-binding specificity of GBPs (97). The algorithms used to perform these analyses will be used in developing more sophisticated software to automate the analyses of hundreds of glycans on SGMs. As glycans on each SGM are structurally defined, they become components of an ever-increasing defined glycan array and continue to expand the collection of structures in the process of defining the human glycome.

*Acknowledgment*—We thank Jamie Heimbürg-Molinaro for critical editing assistance.

\* This work was supported, in whole or in part, by National Institutes of Health Grants GM85448 (to D.F.S.) and GM98791 (to R.D.C.).

‡ To whom correspondence may be addressed: David F. Smith, Dept. of Biochemistry, Emory University School of Medicine, 1510 Clifton Rd. NE, Rm. 4035, Atlanta, GA 30322. Tel.: 404-727-6155; Fax: 404-727-2738; email: dfsmith@emory.edu and Richard D. Cummings, William Patterson Timmie Professor and Chair, Dept. of Biochemistry, Emory University School of Medicine, 1510 Clifton Rd. NE, Ste. 4001, Atlanta, GA 30322. Tel.: 404-727-0988; Fax: 404-727-2738; email: rdcummi@emory.edu.

### REFERENCES

- Mechref, Y., Hu, Y., Garcia, A., and Hussein, A. (2012) Identifying cancer biomarkers by mass spectrometry-based glycomics. *Electrophoresis* **33**, 1755–1767
- North, S. J., Chalabi, S., Sutton-Smith, M., Dell, A., and Haslam, S. M. (2009) in *Handbook of Glycomics* (Cummings, R. D., and Pierce, J. M., eds) Academic Press, San Diego pp. 263–327
- Zoldoš, V., Novokmet, M., Bečeheli, I., and Lauc, G. (2013) Genomics and epigenomics of the human glycome. *Glycoconj. J.* **30**, 41–50
- Cummings, R. D. (2009) The repertoire of glycan determinants in the human glycome. *Mol. Biosyst.* **5**, 1087–1104
- Wuhrer, M. (2013) Glycomics using mass spectrometry. *Glycoconj. J.* **30**, 11–22
- Robinson, L. N., Artpradit, C., Raman, R., Shriver, Z. H., Ruchirawat, M., and Sasisekharan, R. (2012) Harnessing glycomics technologies: integrating structure with function for glycan characterization. *Electrophoresis* **33**, 797–814
- Ito, S. (1969) Structure and function of the glycocalyx. *Fed. Proc.* **28**, 12–25
- Lee, G. M., Johnstone, B., Jacobson, K., and Caterson, B. (1993) The dynamic structure of the pericellular matrix on living cells. *J. Cell Biol.* **123**, 1899–1907
- Weinbaum, S., Tarbell, J. M., and Damiano, E. R. (2007) The structure and function of the endothelial glycocalyx layer. *Annu. Rev. Biomed. Eng.* **9**, 121–167
- Schachter, H., and Freeze, H. H. (2009) Glycosylation diseases: quo vadis? *Biochim. Biophys. Acta* **1792**, 925–930
- Narimatsu, H., Sawaki, H., Kuno, A., Kaji, H., Ito, H., and Ikehara, Y. (2010) A strategy for discovery of cancer glyco-biomarkers in serum using newly developed technologies for glycoproteomics. *FEBS J.* **277**, 95–105
- Nairn, A. V., and Moremen, K. W. (2009) in *Handbook of Glycomics* (Cummings, R. D., and Pierce, J. M., eds) pp. 95–136, Academic Press, San Diego
- Hollenberg, M. D., Fishman, P. H., Bennett, V., and Cuatrecasas, P. (1974) Cholera toxin and cell growth: role of membrane gangliosides. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4224–4228
- Varki, A., and Gagneux, P. (2012) Multifarious roles of sialic acids in immunity. *Ann. N.Y. Acad. Sci.* **1253**, 16–36
- Paulson, J. C., Sadler, J. E., and Hill, R. L. (1979) Restoration of specific myxovirus receptors to asialoerythrocytes by incorporation of sialic acid with pure sialyltransferases. *J. Biol. Chem.* **254**, 2120–2124
- Imberty, A., and Varrot, A. (2008) Microbial recognition of human cell surface glycoconjugates. *Curr. Opin. Struct. Biol.* **18**, 567–576
- Viswanathan, K., Chandrasekaran, A., Srinivasan, A., Raman, R., Sasisekharan, V., and Sasisekharan, R. (2010) Glycans as receptors for influenza pathogenesis. *Glycoconj. J.* **27**, 561–570

18. Burger, M. M. (1969) A difference in the architecture of the surface membrane of normal and virally transformed cells. *Proc. Natl. Acad. Sci. U.S.A.* **62**, 994–1001
19. Hakomori, S. (1984) Tumor-associated carbohydrate antigens. *Annu. Rev. Immunol.* **2**, 103–126
20. Raman, R., Raguram, S., Venkataraman, G., Paulson, J. C., and Sasisekharan, R. (2005) Glycomics: an integrated systems approach to structure-function relationships of glycans. *Nat. Methods* **2**, 817–824
21. Hirabayashi, J., Arata, Y., and Kasai, K. (2001) Glycome project: concept, strategy, and preliminary application to *Caenorhabditis elegans*. *Proteomics* **1**, 295–303
22. Gabius, H. J. (2011) Glycobiomarkers by glycoproteomics and glycan profiling (glycomics): emergence of functionality. *Biochem. Soc. Trans.* **39**, 399–405
23. Doherty, M., McManus, C. A., Duke, R., and Rudd, P. M. (2012) High throughput quantitative N-glycan analysis of glycoproteins. *Methods Mol. Biol.* **899**, 293–313
24. Rakus, J. F., and Mahal, L. K. (2011) New technologies for glycomic analysis: toward a systematic understanding of the glycome. *Annu. Rev. Anal. Chem.* **4**, 367–392
25. Esposito, D., Hurevich, M., Castagner, B., Wang, C. C., and Seeberger, P. H. (2012) Automated synthesis of sialylated oligosaccharides. *Beilstein J. Org. Chem.* **8**, 1601–1609
26. Hsu, C. H., Hung, S. C., Wu, C. Y., and Wong, C. H. (2011) Toward automated oligosaccharide synthesis. *Angew. Chem. Int. Ed. Engl.* **50**, 11872–11923
27. North, S. J., Hitchen, P. G., Haslam, S. M., and Dell, A. (2009) Mass spectrometry in the analysis of N-linked and O-linked glycans. *Curr. Opin. Struct. Biol.* **19**, 498–506
28. Leymarie, N., and Zaia, J. (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal. Chem.* **84**, 3040–3048
29. Nairn, A. V., Aoki, K., dela Rosa, M., Porterfield, M., Lim, J. M., Kulik, M., Pierce, J. M., Wells, L., Dalton, S., Tiemeyer, M., and Moremen, K. W. (2012) Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis. *J. Biol. Chem.* **287**, 37835–37856
30. Antonopoulos, A., North, S. J., Haslam, S. M., and Dell, A. (2011) Glycosylation of mouse and human immune cells: insights emerging from N-glycomics analyses. *Biochem. Soc. Trans.* **39**, 1334–1340
31. Goldberg, D., Sutton-Smith, M., Paulson, J., and Dell, A. (2005) Automatic annotation of matrix-assisted laser desorption/ionization N-glycan spectra. *Proteomics* **5**, 865–875
32. Pang, P. C., Chiu, P. C., Lee, C. L., Chang, L. Y., Panico, M., Morris, H. R., Haslam, S. M., Khoo, K. H., Clark, G. F., Yeung, W. S., and Dell, A. (2011) Human sperm binding is mediated by the sialyl-Lewis(x) oligosaccharide on the zona pellucida. *Science* **333**, 1761–1764
33. Kannagi, R. (2007) Carbohydrate antigen sialyl Lewis a—its pathophysiological significance and induction mechanism in cancer progression. *Chang Gung Med. J.* **30**, 189–209
34. Arend, P. (2011) “Natural” antibodies and histo-blood groups in biological development with respect to histo-blood group A. A perspective review. *Immunobiology* **216**, 1318–1321
35. Miyagawa, S., Ueno, T., Nagashima, H., Takama, Y., and Fukuzawa, M. (2012) Carbohydrate antigens. *Curr. Opin. Organ Transplant.* **17**, 174–179
36. Merkle, R. K., and Cummings, R. D. (1987) Lectin affinity chromatography of glycopeptides. *Methods Enzymol.* **138**, 232–259
37. Cummings, R. D. (1994) Use of lectins in analysis of glycoconjugates. *Methods Enzymol.* **230**, 66–86
38. Hirabayashi, J., Kuno, A., and Tateno, H. (2011) Lectin-based structural glycomics: a practical approach to complex glycans. *Electrophoresis* **32**, 1118–1128
39. Mahal, L. K. (2008) Glycomics: toward bioinformatic approaches to understanding glycosylation. *Anticancer Agents Med. Chem.* **8**, 37–51
40. Fanayan, S., Hincapie, M., and Hancock, W. S. (2012) Using lectins to harvest the plasma/serum glycoproteome. *Electrophoresis* **33**, 1746–1754
41. Patnaik, S. K., and Stanley, P. (2006) Lectin-resistant CHO glycosylation mutants. *Methods Enzymol.* **416**, 159–182
42. Novotny, M. V., Alley, W. R., Jr., and Mann, B. F. (2013) Analytical glycobiochemistry at high sensitivity: current approaches and directions. *Glycoconj. J.* **30**, 89–117
43. Kato, K. (2012) Structural glycomics approaches to molecular recognition events on cell surfaces. *Adv. Exp. Med. Biol.* **749**, 15–32
44. Ruhaak, L. R., and Lebrilla, C. B. (2012) Advances in analysis of human milk oligosaccharides. *Adv. Nutr.* **3**, 406S–414S
45. Fuchs, B. (2012) Analysis of phospholipids and glycolipids by thin-layer chromatography-matrix-assisted laser desorption and ionization mass spectrometry. *J. Chromatogr. A* **1259**, 62–73
46. van Die, I., van Stijn, C. M., Geyer, H., and Geyer, R. (2010) Structural and functional analysis of glycosphingolipids of *Schistosoma mansoni*. *Methods Enzymol.* **480**, 117–140
47. Song, X., Smith, D. F., and Cummings, R. D. (2012) Nonenzymatic release of free reducing glycans from glycosphingolipids. *Anal. Biochem.* **429**, 82–87
48. Huang, Y., Mechref, Y., and Novotny, M. V. (2001) Microscale nonreductive release of O-linked glycans for subsequent analysis through MALDI mass spectrometry and capillary electrophoresis. *Anal. Chem.* **73**, 6063–6069
49. Bigge, J. C., Patel, T. P., Bruce, J. A., Goulding, P. N., Charles, S. M., and Parekh, R. B. (1995) Nonselective and efficient fluorescent labeling of glycans using 2-aminobenzamide and anthranilic acid. *Anal. Biochem.* **230**, 229–238
50. Song, X., Xia, B., Stowell, S.R., Lasanajak, Y., Smith, D.F., Cummings, R.D. (2009) Novel fluorescent glycan microarray strategy reveals ligands for galectins. *Chem. Biol.* **16**, 36–47
51. Townsend, R. R., Lipniunas, P. H., Bigge, C., Ventom, A., and Parekh, R. (1996) Multimode high-performance liquid chromatography of fluorescently labeled oligosaccharides from glycoproteins. *Anal. Biochem.* **239**, 200–207
52. Harvey, D. J. (2012) Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update for 2007–2008. *Mass Spectrom. Rev.* **31**, 183–311
53. Jiménez-Barbero, J., Díaz, M. D., and Nieto, P. M. (2008) NMR structural studies of oligosaccharides related to cancer processes. *Anticancer Agents Med. Chem.* **8**, 52–63
54. Imberty, A. (1997) Oligosaccharide structures: theory versus experiment. *Curr. Opin. Struct. Biol.* **7**, 617–623
55. Kondakova, A. N., Novototskaya-Vlasova, K. A., Arbatsky, N. P., Drutskaya, M. S., Shcherbakova, V. A., Shashkov, A. S., Gilichinsky, D. A., Nedospasov, S. A., and Knirel, Y. A. (2012) Structure of the O-specific polysaccharide from the lipopolysaccharide of *Psychrobacter cryohalolentis* K5(T) containing a 2,3,4-triacetamido-2,3,4-trideoxy-L-arabinose moiety. *J. Nat. Prod.* **75**, 2236–2240
56. Bones, J., Mittermayr, S., O'Donoghue, N., Guttman, A., and Rudd, P. M. (2010) Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer-associated alterations in glycosylation. *Anal. Chem.* **82**, 10208–10215
57. Gillmeister, M. P., Tomiya, N., Jacobia, S. J., Lee, Y. C., Gorfien, S. F., and Betenbaugh, M. J. (2009) An HPLC-MALDI MS method for N-glycan analyses using smaller size samples: application to monitor glycan modulation by medium conditions. *Glycoconj. J.* **26**, 1135–1149
58. Goodarzi, M. T., and Turner, G. A. (1998) Reproducible and sensitive determination of charged oligosaccharides from haptoglobin by PNGase F digestion and HPAEC/PAD analysis: glycan composition varies with disease. *Glycoconj. J.* **15**, 469–475
59. Rohrer, J. S., and Townsend, R. R. (1995) Separation of partially desialylated branched oligosaccharide isomers containing  $\alpha(2\rightarrow3)$ - and  $\alpha(2\rightarrow6)$ -linked Neu5Ac. *Glycobiology* **5**, 391–395
60. Kawar, Z. S., Johnson, T. K., Natunen, S., Lowe, J. B., and Cummings, R. D. (2008) PSGL-1 from the murine leukocytic cell line WEHI-3 is enriched for core 2-based O-glycans with sialyl Lewis x antigen. *Glycobiology* **18**, 441–446
61. Wuhrer, M., Koeleman, C. A., and Deelder, A. M. (2009) Two-dimensional HPLC separation with reverse-phase-nano-LC-MS/MS for the characterization of glycan pools after labeling with 2-aminobenzamide. *Methods Mol. Biol.* **534**, 79–91
62. Tomiya, N., Awaya, J., Kurono, M., Endo, S., Arata, Y., and Takahashi, N. (1988) Analyses of N-linked oligosaccharides using a two-dimensional mapping technique. *Anal. Biochem.* **171**, 73–90
63. Hase, S., Ikenaka, K., Mikoshiba, K., and Ikenaka, T. (1988) Analysis of tissue glycoprotein sugar chains by two-dimensional high-performance

- liquid chromatographic mapping. *J. Chromatogr.* **434**, 51–60
64. Volpi, N., and Maccari, F. (2006) Electrophoretic approaches to the analysis of complex polysaccharides. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **834**, 1–13
  65. Jackson, P. (1996) The analysis of fluorophore-labeled carbohydrates by polyacrylamide gel electrophoresis. *Mol. Biotechnol.* **5**, 101–123
  66. Arnold, J. N., Saldova, R., Galligan, M. C., Murphy, T. B., Mimura-Kimura, Y., Telford, J. E., Godwin, A. K., and Rudd, P. M. (2011) Novel glycan biomarkers for the detection of lung cancer. *J. Proteome Res.* **10**, 1755–1764
  67. Cummings, R. D., and Kornfeld, S. (1982) Fractionation of asparagine-linked oligosaccharides by serial lectin-agarose affinity chromatography. A rapid, sensitive, and specific technique. *J. Biol. Chem.* **257**, 11235–11240
  68. Osawa, T., and Tsuji, T. (1987) Fractionation and structural assessment of oligosaccharides and glycopeptides by use of immobilized lectins. *Annu. Rev. Biochem.* **56**, 21–42
  69. Diaz, S., and Varki, A. (2001) Metabolic radiolabeling of animal cell glycoconjugates. *Curr. Protoc. Mol. Biol.* Chapter 17, Unit 17.14
  70. Yamashita, K., Mizuochi, T., and Kobata, A. (1982) Analysis of oligosaccharides by gel filtration. *Methods Enzymol.* **83**, 105–126
  71. Royle, L., Radcliffe, C. M., Dwek, R. A., and Rudd, P. M. (2006) Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods Mol. Biol.* **347**, 125–143
  72. Varki, A., and Kornfeld, S. (1980) Structural studies of phosphorylated high mannose-type oligosaccharides. *J. Biol. Chem.* **255**, 10847–10858
  73. Nyame, K., Cummings, R. D., and Damian, R. T. (1988) Characterization of the high mannose asparagine-linked oligosaccharides synthesized by *Schistosoma mansoni* adult male worms. *Mol. Biochem. Parasitol.* **28**, 265–274
  74. Cummings, R. D., and Kornfeld, S. (1982) Characterization of the structural determinants required for the high affinity interaction of asparagine-linked oligosaccharides with immobilized *Phaseolus vulgaris* leucoagglutinating and erythroagglutinating lectins. *J. Biol. Chem.* **257**, 11230–11234
  75. Cummings, R. D., Trowbridge, I. S., and Kornfeld, S. (1982) A mouse lymphoma cell line resistant to the leucoagglutinating lectin from *Phaseolus vulgaris* is deficient in UDP-GlcNAc:α-D-mannoside β1,6 N-acetylglucosaminyltransferase. *J. Biol. Chem.* **257**, 13421–13427
  76. Granovsky, M., Fata, J., Pawling, J., Muller, W. J., Khokha, R., and Dennis, J. W. (2000) Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nat. Med.* **6**, 306–312
  77. Ahn, Y. H., Lee, J. Y., Lee, J. Y., Kim, Y. S., Ko, J. H., and Yoo, J. S. (2009) Quantitative analysis of an aberrant glycoform of TIMP1 from colon cancer serum by L-PHA-enrichment and SISCAPA with MRM mass spectrometry. *J. Proteome Res.* **8**, 4216–4224
  78. Stowell, S. R., Arthur, C. M., Mehta, P., Slanina, K. A., Blixt, O., Leffler, H., Smith, D. F., and Cummings, R. D. (2008) Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens. *J. Biol. Chem.* **283**, 10109–10123
  79. Stowell, S. R., Arthur, C. M., Dias-Baruffi, M., Rodrigues, L. C., Gourdine, J. P., Heimbürg-Molinari, J., Ju, T., Molinari, R. J., Rivera-Marrero, C., Xia, B., Smith, D. F., and Cummings, R. D. (2010) Innate immune lectins kill bacteria expressing blood group antigen. *Nat. Med.* **16**, 295–301
  80. Xia, B., Kowar, Z. S., Ju, T., Alvarez, R. A., Sachdev, G. P., and Cummings, R. D. (2005) Versatile fluorescent derivatization of glycans for glycomic analysis. *Nat. Methods* **2**, 845–850
  81. de Boer, A. R., Hokke, C. H., Deelder, A. M., and Wührer, M. (2007) General microarray technique for immobilization and screening of natural glycans. *Anal. Chem.* **79**, 8107–8113
  82. Song, X., Lasanajak, Y., Xia, B., Heimbürg-Molinari, J., Rhea, J. M., Ju, H., Zhao, C., Molinari, R. J., Cummings, R. D., and Smith, D. F. (2011) Shotgun glycomics: a microarray strategy for functional glycomics. *Nat. Methods* **8**, 85–90
  83. Kornfeld, R., and Ferris, C. (1975) Interaction of immunoglobulin glycopeptides with concanavalin A. *J. Biol. Chem.* **250**, 2614–2619
  84. Baenziger, J. U., and Fiete, D. (1979) Structural determinants of concanavalin A specificity for oligosaccharides. *J. Biol. Chem.* **254**, 2400–2407
  85. Shibuya, N., Goldstein, I. J., Broekaert, W. F., Nsimba-Lubaki, M., Peeters, B., and Peumans, W. J. (1987) Fractionation of sialylated oligosaccharides, glycopeptides, and glycoproteins on immobilized elderberry (*Sambucus nigra* L.) bark lectin. *Arch. Biochem. Biophys.* **254**, 1–8
  86. Shibuya, N., Goldstein, I. J., Broekaert, W. F., Nsimba-Lubaki, M., Peeters, B., and Peumans, W. J. (1987) The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(α2-6)Gal/GalNAc sequence. *J. Biol. Chem.* **262**, 1596–1601
  87. Smith, D. F., Song, X., and Cummings, R. D. (2010) Use of glycan microarrays to explore specificity of glycan-binding proteins. *Methods Enzymol.* **480**, 417–444
  88. Wang, W. C., and Cummings, R. D. (1988) The immobilized leucoagglutinin from the seeds of *Maackia amurensis* binds with high affinity to complex-type Asn-linked oligosaccharides containing terminal sialic acid-linked α-2,3 to penultimate galactose residues. *J. Biol. Chem.* **263**, 4576–4585
  89. Itakura, Y., Nakamura-Tsuruta, S., Kominami, J., Sharon, N., Kasai, K., and Hirabayashi, J. (2007) Systematic comparison of oligosaccharide specificity of *Ricinus communis* agglutinin I and *Erythrina* lectins: a search by frontal affinity chromatography. *J. Biochem.* **142**, 459–469
  90. Kochibe, N., and Furukawa, K. (1980) Purification and properties of a novel fucose-specific hemagglutinin of *Aleuria aurantia*. *Biochemistry* **19**, 2841–2846
  91. Matsumoto, I., and Osawa, T. (1969) Purification and characterization of an anti-H(O) phytohemagglutinin of *Ulex europaeus*. *Biochim. Biophys. Acta* **194**, 180–189
  92. Yu, Y., Mishra, S., Song, X., Lasanajak, Y., Bradley, K. C., Tappert, M. M., Air, G. M., Steinhauer, D. A., Halder, S., Cotmore, S., Tattersall, P., Agbandje-McKenna, M., Cummings, R. D., and Smith, D. F. (2012) Functional glycomic analysis of human milk glycans reveals the presence of virus receptors and embryonic stem cell biomarkers. *J. Biol. Chem.* **287**, 44784–44799
  93. Niñonuevo, M. R., and Lebrilla, C. B. (2009) Mass spectrometric methods for analysis of oligosaccharides in human milk. *Nutr. Rev.* **67**, S216–S226
  94. Natunen, S., Satomaa, T., Pitkänen, V., Salo, H., Mikkola, M., Natunen, J., Otonkoski, T., and Valmu, L. (2011) The binding specificity of the marker antibodies Tra-1–60 and Tra-1–81 reveals a novel pluripotency-associated type 1 lactosamine epitope. *Glycobiology* **21**, 1125–1130
  95. Wu, S., Tao, N., German, J. B., Grimm, R., and Lebrilla, C. B. (2010) Development of an annotated library of neutral human milk oligosaccharides. *J. Proteome Res.* **9**, 4138–4151
  96. Kobata, A. (2010) Structures and application of oligosaccharides in human milk. *Proc. Jpn. Acad. Ser. B, Phys. Biol. Sci.* **86**, 731–747
  97. Cholleti, S. R., Agravat, S., Morris, T., Saltz, J. H., Song, X., Cummings, R. D., and Smith, D. F. (2012) Automated motif discovery from glycan array data. *Omics* **16**, 497–512