Functional Glycomic Analysis of Human Milk Glycans Reveals Presence of Virus Receptors and Embryonic Stem Cell Biomarkers

Ying Yu1,2, Shreya Mishra1, Xuezheng Song1, Yi Lasanajak1, Konrad C. Bradley2, Mary M. Tappert3, Gillian M. Air3, David A. Steinhauer2, Sujata Halder4, Susan Cotmore5, Peter Tattersall5, Mavis Agbandje-McKenna4, Richard D. Cummings1 and David F. Smith1

1Department of Biochemistry and the Glycomics Center; Emory University School of Medicine; Atlanta, GA 30322; 2Department of Microbiology and Immunology; Emory University School of Medicine; Atlanta, GA 30322; 3Department of Biochemistry & Molecular Biology; University of Oklahoma Health Sciences Center; Oklahoma City, OK 73126; 4Department of Biochemistry & Molecular Biology; University of Florida, Gainesville, FL 32607; 5Department of Laboratory Medicine, Yale School of Medicine, New Haven, CT 06519.

Running Title: Functional Glycomic Analysis of Human Milk Glycans

Send Correspondence to: David F. Smith, Department of Biochemistry, Emory University School of Medicine, O. Wayne Rollins Research Center, 1510 Clifton Road, Suite 4001, Atlanta, GA 30322, Tel: 404-727-6155; Fax: 404-727-2738; Email: dfsmith@emory.edu and Richard D. Cummings, Department of Biochemistry, Emory University School of Medicine, O. Wayne Rollins Research Center, 1510 Clifton Road, Suite 4001, Atlanta, GA 30322, Tel: 404-727-5962 (main office); Fax: 404-727-2738; Email: rdcummi@emory.edu

Keywords: Human milk glycans, glycan microarray, functional glycomics, virus interaction, biomarkers for human embryonic stem cells

Background: Recognition of human milk glycans (HMGs) by lectins, antibodies, and pathogens is poorly understood.

Results: Microarrays of isolated HMGs exhibited specific binding to proteins and pathogens.

Conclusion: HMG microarrays interrogation and novel metadata-assisted glycan sequencing provide a functional glycomics approach to discovering HMG function.

Significance: HMGs represent a potential “liquid innate immune system” that is specifically recognized by antibodies and pathogens.

SUMMARY

Human milk contains a large diversity of free glycans beyond lactose, but their functions are not well understood. To explore their functional recognition, here we describe a shotgun glycan microarray prepared from isolated human milk glycans (HMGs), and our studies on their recognition by viruses, antibodies, and glycan-binding proteins (GBPs), including lectins. The total neutral and sialylated HMGs were derivatized with a bifunctional fluorescent tag, separated by multidimensional HPLC, and archived in a tagged glycan library, which was then used to print a shotgun glycan microarray (SGM). This SGM was first interrogated with well-defined GBPs and antibodies. These data demonstrated both the utility of the array, and provided preliminary structural information (metadata) about this complex glycome. Anti-TRA-1 antibodies that recognize human pluripotent stem cells specifically recognized several HMGs, which were then further structurally defined as novel epitopes for these antibodies. Human influenza viruses and Parvovirus Minute Viruses of Mice also specifically recognized several HMGs. For glycan sequencing, we used a novel approach termed Metadata-Assisted Glycan Sequencing (MAGS), in which we combine information from analyses of glycans by mass spectrometry with glycan
interactions with defined GBPs and antibodies before and after exoglucosidase treatments on the microarray. Together, these results provide novel insights into diverse recognition functions of HMGs and show the utility of the SGM approach and MAGS as resources for defining novel glycan recognition by GBPs, antibodies, and pathogens.

INTRODUCTION

As the natural food source for newborns, human milk provides not only all of the nutrients necessary for infants to grow and develop, but also provides health benefits in early childhood (1). In addition to antibodies that greatly enhance defense of the infant against various diseases, human milk possesses a rich pool of free-reducing oligosaccharides (glycans) that generally are unique in composition to humans and differ from those in the milk of other mammals (2,3). The mature milk is estimated to contain 5-15 g/L free human milk glycans (HMGs), depending on the individual’s lactation time and blood group type (4,5). So far, >200 unique HMGs have been detected and >100 have been structurally characterized (6,7). The large number and the remarkable structural diversity of these glycans suggest that they possess multiple biological effects (7-9). The prebiotic and anti-adhesion effects are the most common functions attributed to HMGs (10,11). Furthermore, HMGs are associated with neonatal intestinal development (12-14) and limiting risks of necrotizing enterocolitis (7).

Although the multiple functions of HMGs are poorly understood, numerous studies have shown that their functional activities are dependent on their structures. For example, sialylated HMGs inhibit cholera toxin binding (15) and leukocyte adhesion to cultured human umbilical vein endothelial cells (16). Specific fucosylated HMGs are recognized by enteropathogens, including Helicobacter pylori (17), rabbit calicivirus (18) and Norwalk virus (19). Neutral HMGs, especially H type-2 glycans, inhibit Campylobacter jejuni adherence to Hep-2 cells and intestinal mucosa (20). Although the reported in vitro and in vivo data provide important information for understanding the effect of HMGs, typical experiments with HMG utilize either a small number of defined glycans or mixtures of HMG fractions. Such limitations represent challenges in studying HMGs, where the goal is to determine the roles of specific glycans in the milk glycome and to establish the relationships between glycan structures and their biological effects. However, linking function to structures of HMGs is difficult; many HMGs are comprised of linear and branched polymers of type 1 and type 2 lactosamine, Galβ1-3GlcNAc and Galβ1-4GlcNAc, respectively, substituted with α-linked Neu5Ac and Fuc. It is difficult to assign structures by mass spectrometry alone because of isobaric and isomeric structures, and a wide variety of approaches are often required, thus hindering progress in this area (21,22).

Others and we have made extensive use of glycan microarrays with defined chemo-enzymatically derived glycans to explore glycan recognition by glycan-binding proteins (GBPs) and microorganisms (23-27). However, because glycan synthesis is difficult, only a small fraction of the predicted, large number of glycans in the human glycome (28) are available for array production. To address this limitation we developed an alternative strategy termed “Shotgun Glycomics” (29) in which mixtures of free glycans derived from glycoproteins and glycolipids are derivatized with a bifunctional fluorescent tag, separated by multidimensional HPLC, and individual glycans are printed as a shotgun glycan microarray (SGM). In this approach glycan structures are defined after they are identified through their recognition by a GBP or pathogen, and therefore potentially functionally important. Here we have applied this approach to HMGs, and defined those HMGs that are individually recognized by selected antibodies and pathogens. In addition, we combined the use of mass spectrometry, recognition by defined GBPs, and exoglucosidase treatments to help provide more detailed information about specific glycan structures in an approach termed Metadata-Assisted Glycan Sequencing (MAGS). This work represents the first use of a shotgun...
glycomics approach to prepare a natural glycan microarray of HMG containing >100 glycans.

EXPERIMENTAL PROCEDURES

Materials—Free reducing glycans used as standards were purchased from Sigma-Aldrich (St. Louis, MO) and V-LABS, Inc. (Covington, LA). All standard chemicals were bought from Sigma-Aldrich and used without further purification. Human milk was purchased from the Mothers Milk Bank (Austin, TX). Asialo, biantennary N-glycan (NA2) was prepared by mild acid hydrolysis and PNGase F digestion of a chicken egg yolk glycopeptide, prepared as described (30). 2-Amino-N-(2-aminoethyl)benzamide (AEAB) was synthesized as described previously (31). β1-3 galactosidase, β1-4 galactosidase, and α2-3 neuraminidase were obtained from New England BioLabs (Ipswich, MA); neuraminidase (Arthrobacter ureafaciens) was from Roche Applied Science (Indianapolis, IN); Jack bean β-galactosidase was from ProZyme (Hayward, CA). HPLC solvents were purchased from ThermoFisher Scientific (Waltham, MA). An Ultraflex-II TOF/TOF system from Bruker Daltonics was used for MALDI-TOF mass spectrometry analysis of glycan conjugates. Biotinylated lectins were from Vector Labs (Burlingame CA), Cy5- and Alexa488-labeled Streptavidin were from Invitrogen Life Technologies (Grand Island, NY). Anti-blood group Lewis a (Lea) and anti-CD15 antibodies were from Santa Cruz Biotechnology Inc. (Santa Cruz, CA). Mouse monoclonal antibodies to blood group H1 antigen (ab3355) and to Lewis b (Leb) (ab3968) were purchased from Abcam (Cambridge, MA). Anti-Sialyl Leα antibody was a kind gift from Professor Nancy Louis (Emory University School of Medicine, Atlanta, GA) and the anti-type-1 chain antibody was a kind gift from Dr. Irving Weissman (Stanford University, Stanford, CA). Anti-Leα antibody was prepared in the Cummings lab (Mandalasi et al., manuscript under review at Glycobiology). Anti-TRA-1-81 and TRA-1-60 monoclonal antibodies were purchased from Millipore (Billerica, MA). Printing of glycan arrays was performed using a Piezorray Printer (Perkin Elmer) on NEXTERION® Slide H slides, which are NHS-activated slides from Schott Nexterion (Schott North America, Louisville, KY), as previously described (29,32).

Preparation of human milk glycans — The procedure for isolation of HMGs is illustrated in Fig. S1A (33). Briefly, human milk was defatted by centrifugation at 6,000xg for 30 minutes (4°C); skimmed milk was filtered through glass wool and mixed with 2 volumes of ethanol and allowed to stand at 4°C overnight to precipitate the bulk of the lactose and proteins. After centrifugation, the supernatant was concentrated and fractionated with Sephadex G-25 column to fraction A, B and C. Fraction A, enriched with glycans larger than lactose, was applied to a DEAE column equilibrated with 2 mM pyridinium acetate, yielding neutral, monosialyl, and disialyl fractions by eluting sequentially with 2 mM, 20 mM, and 200 mM pyridine acetate, respectively (34). The resulting fractions were lyophilized.

Glycan-AEAB conjugation and purification — Standard glycans and the 3 human milk glycan fractions were conjugated with AEAB as described previously (31). Briefly, 1-10 mg glycan was mixed with 50-250 µL of freshly prepared 0.35 M AEAB hydrochloride salt and equal volume of 1 M NaCNBH4 in DMSO/AcOH (v/v = 7/3). The conjugation reaction was left to proceed for 2 hours at 65°C and stopped by addition of 10 volumes of cold acetonitrile and allowed to stand for 30 min at -20°C. The precipitated Glycan-AEAB derivatives (GAEABs) were collected after centrifugation at 10,000x g for 3 minutes.

High performance liquid chromatography — A Shimadzu HPLC CBM-20A system, coupled with a UV detector SPD-20A and a fluorescence detector RF-10Axl, was used for HPLC analysis and separation of GAEABs. The AEAB-conjugated milk glycans were first separated by normal phase HPLC on an Agilent NH2 column (250 x 4.6 mm). The mobile phase linear gradient was from 80% acetonitrile, 10 mM ammonium acetate, pH 4.5 to 10% acetonitrile, 125 mM ammonium acetate, pH 4.5 in 160-200 minutes. Individual peaks were collected and dried. Each peak collected on the NH2 column
was further purified in a second dimension by reverse phase HPLC on a porous graphitized carbon (PGC) column (150 x 4.6 mm) with a gradient of 15% to 45% acetonitrile (0.1% trifluoroacetic acid, TFA) in 30 minutes or 15 minutes. The effluents were monitored by UV absorption at 330 nm and/or fluorescence at 420 nm with excitation 330 nm. LNFPIII–AEAB or lactose–AEAB was utilized as the standard for the quantification of the AEAB derivatives. The individual glycans after the second dimension chromatography were quantified, dried, and reconstituted in water at a concentration of 200 µM and stored as a HMG tagged glycan library (TGL).

Printing of the SGM, binding assays, and analysis of microarray—The printing conditions for the SGM, binding assays, and data analysis of glycan microarrays have been previously described (31,35,36). Briefly, each glycan in the TGL was adjusted to 100 µM in sodium phosphate buffer (100 mM, pH 8.5) in a final volume of 10 µl and distributed into the 384-well source plate of the Piezorray printer. The glycans were printed as 0.3 nL aliquots in replicates of 5. The biotinylated lectins or unlabeled antibodies were incubated for 1-3 hours on the slides. After washing, the bound lectins were detected by a secondary incubation with cyanine 5–labeled streptavidin (Invitrogen) (5 µg/ml) and the bound antibodies were detected using Alexa488–labeled goat anti-mouse IgG (Invitrogen) (5 µg/ml). The fluorescence signals were generated using a ProScanArray Scanner (Perkin Elmer - Waltham, MA) with the excitation/emission wavelength set at 649/670 nm and 495/510 nm for cyanine 5 and Alexa488, respectively.

Virus preparations and binding assays — The human parainfluenza (hPIV) and influenza A were grown in LLC-MK2 cells and MDCK cells, respectively, as previously described (37,38). Harvested viruses were purified through a sucrose gradient centrifugation (10-60% for hPIV and 10-40% for influenza A) and resuspended in calcium-magnesium saline, followed by labeling with Alexa488 succinimidyl ester (Molecular Probes, Invitrogen, Life Technologies, Grand Island, NY). For the binding assays, the slides were incubated with labeled viruses for 1 hour at 4°C and directly scanned after washing. The WT and mutant strains of minute virus of mice (MVM) were constructed and prepared as previously described (39-42). Briefly, the WT MVMp empty particles were grown in Mouse A9 ouab’11 fibroblasts, MVMi mutants (MVMi-ggA and MVMi-agD) were grown in NB324K cells and VLP (virus-like particle) were grown in sf9 insect cells. The harvested viruses were released by repeated freezing and thawing, and then purified by centrifugal sedimentation and density equilibrium gradients. The purity and integrity of prepared viruses were assessed by hemagglutination, SDS-PAGE, and electron microscopy (41).

Exoglycosidase digestion of glycans printed on a SGM — On-array exoglycosidase digestions were carried out at 37°C in buffers recommended by the suppliers of the enzymes. Enzymes were used without further purification. Microarrays were rehydrated in the digestion buffer for 5 min prior to enzyme addition and washed 4 times with TSM buffer (20 mM Tris-HCl, 150 mM sodium chloride, 2 mM calcium chloride and 2 mM magnesium chloride) supplemented with 0.1% Tween-20, 4 times with TSM buffer, and 4 times with water before detection with lectins or antibodies. Reaction conditions were optimized for each exoglycosidase.

RESULTS

Preparation of the human milk glycome SGM — The generation of a SGM of HMGs from a single individual milk sample is illustrated in Fig. 1. To extract the human milk free glycome, the lipids, proteins and most of the lactose were removed by centrifugation and ethanol precipitation. The glycans were obtained after size exclusion chromatography and then separated into three distinct groups (neutral, monosialyl and disialyl) by ion-exchange chromatography (Fig. S1) (33). Adopting our shotgun glycomics approach, the glycan mixtures in each group were conjugated with AEAB and separated by two-dimensional HPLC into a total of 156 individual fractions. Selected
fractions were subjected to an additional round of reverse-phase HPLC to further improve the purity and obtain 127 glycans in the TGL. The array contains 73 (57%) sialylated glycans and 54 (43%) neutral glycans. It should be noted that although the number of sialylated glycans exceeds the neutral ones, the total abundance of neutral glycans is much higher than that of sialylated ones. Analysis of the observed masses of all samples (Table S1) revealed glycans ranging in size from 2 residues (2 Hexoses, Hex), i.e., lactose, up to 12 residues (6 Hex + 4 HexNAc + 2 Fuc/Neu5Ac) and in mass from 506.337 [M+H]+ (2 Hex) to 2550 [M+H]+ (6 Hex + 4 HexNAc + 2 Neu5Ac). Of the 54 neutral glycans, over 80% carry one to three fucose residues, while less than half of the sialylated glycans are fucosylated. Of the 73 sialylated glycans, 17 were purified from the monosialyl group and 56 from the disialyl group. However, mass results indicated that some of the samples purified from the disialyl group also carry one sialic acid. After quantification using AEAB fluorescence, the 127 glycans were adjusted to the same concentration (regardless of their abundance) and printed (replicates of n=5) on NHS-activated glass slides along with 11 structurally defined glycans that serve as controls for binding experiments (Table S1). Subsequently, the HMG microarray was interrogated with lectins and antibodies to evaluate whether the glycans were effectively printed.

Preliminary characterization of the HMGs on the SGM by lectins and antibodies – Defined glycan microarrays are used to explore the specificity of GBPs, including lectins and antibodies (23). In turn, GBPs with well-defined binding specificity can assist in the elucidation of glycan structures. To interrogate the SGM and evaluate the structural diversity of isolated HMGs, ten biotinylated lectins, each at several concentrations in the range of 0.001 - 10 µg/ml, were applied to the array. No significant binding was observed with Concanavalin A (Con A), *Vicia villosa* lectin (VVL), *Griffonia simplicifolia* lectin II (GSL-II) and *Maackia amurensis* Lectin I (MAL-I). These lectins recognize glycans containing mannose (43-45), terminal GalNAc (46), terminal GlcNAc (47) and terminal Neu5Acα2-3Galβ1-4GlcNAc (48), respectively; the absence of their binding is consistent with the lack of such structures in human milk (6). Independently the binding of these lectins to a defined glycan microarray from the CFG (v5.0) was evaluated, and the data are available at www.functionalglycomics.org. The other six lectins, *Aleuria aurantia* lectin (AAL), *Sambucus nigra* agglutinin (SNA), *Lotus tetragonolobus* lectin (LTL), *Ulex europaeus* agglutinin I (UEA-I), *Ricinus communis* agglutinin I (RCA-I), and *Erythrina cristagalli* Lectin (ECL) exhibited binding to many of the HMGs on the SGM, as discussed below, and the binding histograms showing average relative fluorescence units (RFU) for selected concentrations of each lectin are shown in Fig. 2, and the binding data are listed in Table S2.

Three fucose-binding lectins, AAL, LTL and UEA-I were used to reveal the fucosylated glycans on the SGM. Approximately 61 glycans (48%) showed strong binding with AAL (RFU>10,000), which binds to terminal α-linked L-fucose in 1-2, 1-3, and 1-4 linkages (49), and an additional 24 glycans (19%) had weaker binding (RFU 1,000 - 8,000). Together, the fucosylated glycans recognized by AAL made up 67% of the total isolated HMGs, which is close to the percentage found by using HPLC-Chip/MS method (21). In addition, most of the neutral glycans and about 40% of the sialylated glycans are fucosylated, in agreement with the results from glycan composition analysis based on mass (Table S1). UEA-I, which is specific for α1-2 linked fucose on a type-2 chain (50,51), bound very weakly (<4,000 RFU) to several multifucosylated, neutral glycans (Fig. 2B) suggesting the presence of α1-2 fucose on our HMG array. However, this weak binding was presumably due to the cross reactivity of UEA-1 with other fucose-containing glycans as shown in Table 1, Fig. 2, and Table S2 and actually not from Fucα1-2. This conclusion was supported by the observation that glycans were resistant to α1-2 fucosidase digestion in solution (data not shown). This result indicated the absence of H type-2 (Fucα1-2Galβ1-4GlcNAc) and Leα (Fucα1-2Galβ1-4(Fucα1-3)GlcNAc) structures in the SGM. LTL recognizes α1-3-linked fucose within type-2 glycans (52,53), like Leα and Leγ.
determinants. It showed strong binding to three neutral glycans (H-75, H-83, and H-87) and weak binding to another five glycans (Fig. 2C). The three high affinity binders likely contain terminal Le\(^a\) determinants because absence of UEA-I binding (Fig. 2B) excludes the possibility of \(\alpha 1-2\)-linked fucose.

SNA, which binds to the Neu5Aca2-6Galβ1-4GlcNAc determinant (23,54), bound well to 14 (out of 73) sialylated glycans on the SGM along with three control glycans (2,6-DS-NA2, LSTc, and fetuin, structures shown in Table S1), indicating the existence of \(\alpha 2,6\)-sialylated type-2 structure in 19% of the glycans (Fig. 2D). Both RCA-I and ECL recognize terminal Galβ1-4GlcNAc (55-59), but the former has much higher affinity and also binds to Neu5Aca2-6Galβ1-4GlcNAc with slightly lower affinity, while the latter can also bind Fucα1-2Galβ1-4Glc (60). Consistent with these features, the binding pattern of RCA-I on this SGM was similar to the combined pattern of SNA and ECL (Fig. 2D-F), together suggesting that ~46% of glycans have terminal type-2 or sialylated type-2 structures. In summary, the results of lectin binding both validated the preparation of the SGM and provided significant structural information on individual HMGs.

The structures of HMGs reflect the Lewis blood type and secretor status of the mothers (61,62). Certain fucosylated glycans that were suggested to possess important biological functions (3) occur only in the milk of Lewis positive and/or Secretor positive mothers due to the expression of FUT-3 and FUT-2, respectively (19). Since there are no lectins specific for Lewis Blood Group antigens, we analyzed the SGM with several blood-group related monoclonal antibodies. As shown in Fig. 3A, the mAb to Le\(^a\) bound to 34 glycans, including 9 sialylated ones, with fluorescence signals in the range of 500-15,000 RFU. To reveal additional Le\(^a\)-containing glycans, we interrogated the array with an anti-sialyl Le\(^a\) (SLe\(^a\)) antibody. According to the data from the structurally defined CFG glycan microarray (http://www.functionalglycomics.org/glycomics/publicdata), this antibody also binds to LSTa (Neu5Aca2-3Galβ1-3GlcNAcβ1-3Galβ1-4Glc) moiety with lower signal. At the concentration of 10 µg/ml, 31 sialylated glycans showed binding signals higher than 30,000 RFU with the antibody, and another 19 glycans had signals over 5,000 RFU (Fig. 3B). Although we could not simply assign SLe\(^a\) and LSTa structures by signal intensity, it is certain that many of these sialyl glycans are fucosylated based on their AAL binding (Fig. 2A) and may contain the SLe\(^a\) moiety. Nevertheless, the abundance of Le\(^a\)-containing glycans showed that the SGM was from a Lewis positive donor. To determine the secretor status of the donor, we interrogated the SGM with anti-Le\(^b\) and anti-blood group H type-1 antibodies. There was little to no binding observed with the anti-Le\(^b\) antibody at any concentration (data not shown), which indicates that the milk sample is from a secretor negative donor. This finding was confirmed by assaying with anti-H1 antibody, which bound only the control LNFP I (Fucα1-2Galβ1-3GlcNAcβ1-6Galβ1-4Glc) (Fig. 3C), the precursor of Le\(^b\) antigen.

We also interrogated the array with anti-CD15 antibody, known to recognize Le\(^x\) antigen. Strong binding was observed to glycans H-98 and H-103 with weaker binding to several other neutral glycans (Fig. 3D). When comparing CD15 antibody with lectin LTL, the two Le\(^x\)-recognizing proteins showed distinct specificity toward different HMGs although with some overlap. This could indicate that the recognition does not solely depend on the Le\(^a\) determinant for complex glycans and the nearby residues or branches could affect the binding.

Detection of HMGs that are epitopes of Anti-TRA-1 antibodies, specific for human pluripotent stem cells – Recently, Natunen et al. (63) predicted epitopes for mAbs anti-TRA-1-60 and anti-TRA-1-81 based on binding data from version 4.2 of the CFG glycan microarray. These mAbs, which are specific for human pluripotent stem cells, bound only to two glycans, both containing the type-1 lactosamine epitope, Galβ1-3GlcNAcβ1-3Galβ1-4GlcNAc, on that version of the CFG glycan microarray. We further examined the two mAbs on Version 5.0 of the CFG array, which contains many multiantennary glycans with poly-N-
acetyllactosamine (64) and observed strong binding by three additional glycans (Fig. S2). Importantly, these three glycans (#522, #572 and #573) are multiantennary glycans with 2 to 3 type-1 lactosamine repeats at their non-reducing ends. In addition, we observed weak but significant binding to 387, a glycan with a fucosylated type-1 lactosamine chain. Considering that HMG is a rich source of type 1 and type 2 lactosamine structures, we interrogated the SGM with the two anti-TRA-1 antibodies at several concentrations (1 µg/ml - 100 µg/ml) and the results at 50 µg/ml are shown in Fig. 3E and 3F. Consistent with the CFG array data, TRA-1-60 and TRA-1-81 share similar receptor specificity since both bind glycans H-71, H-99, H-108, H-111, H-125, and H-127. However, unlike the CFG results, which showed no significant difference in the signal intensity and binding patterns for the two mAbs at 50 µg/ml, we observed that with the SGM the signal intensity for TRA-1-60 binding was always several fold higher than that of TRA-1-81 at the same concentration, and that there are three more low affinity binders for TRA-1-60, H-109, H-112, and H-126 (Fig. 3E). To define the glycan epitope in HMG for TRA-1 antibodies, we retrieved glycans H-71, H-99, H-108, H-111, H-125, and H-127 from the TGL for further characterization. With the exception of sialylated glycan H-71, all of the TRA-1-bound glycans are neutral fucosylated structures consisting of 2-4 lactosamine repeats. We predict from previous studies (63) that these glycans possess type-1 lactosamine. Our data also indicate that the two anti-TRA-1 mAbs recognize complex glycans, as described below.

Investigation of virus binding to the SGM – The HMG-derived SGM provides a library of 127 naturally occurring glycans that permits us to investigate the binding properties of biologically relevant proteins and pathogens, and to provide interesting insights into the potential function of HMGs. To explore the general application of this HMG-derived SGM for exploring pathogen interactions, we examined minute virus of mice (MVM) and influenza virus, both of which attach to the sialic acid on the surface of their target cells at the initial stage of infection (40,65). Two prototype strains, empty capsid (MVMp-WT) and virus-like particle (MVMp-VLP), and two immunosuppressive strain mutants, MVMi-agD and MVMi-ggA (39), were tested at 200 µg/ml concentration and detected by a rabbit anti-MVM capsid antibody. All of the MVM viruses recognized glycans H-30, H-31, H-34, H-43, H-60, and H-73 with each non-WT strain binding several additional glycans (Fig. 4). Interestingly, the initial MALDI data (Table S1) showed that the 6 common binders are disialylated glycans, and the lectin- and antibody-binding data revealed that all of the binders were also recognized by the anti-SLea/LSTa antibody, suggesting that the terminal sialyl α2-3 linked type-1 motif might be part of the binding determinant. This is a new finding compared to the previous report with CFG Glycan Array Version 3.0 (~180 glycans), which concluded that MVMs specifically recognized α2-3 linked type-2 motif and MVMi also bound to α2-8-linked multisialylated glycans (41). These data are complementary to the data from the CFG array since there is very little overlap between the structures found on the CFG array and HMG-derived SGM. In fact, these observations suggest that the repertoire of glycan receptors of MVM is broader than originally reported and these viruses prefer highly charged or possibly multisialylated glycans. We retrieved glycans that were bound by all MVM strains from the TGL for more detailed structural analysis discussed below.

Recently, we utilized the CFG glycan microarray to determine the receptor binding properties of 2009 swine isolates along with some other strains of influenza viruses (37). Our results showed that the highest binding of H1N1 isolates were toward sialylated poly-N-acetyllactosamine structures, which are abundant among HMGs. To determine if HMGs contained natural glycans capable of binding Influenza A virus, we interrogated the SGM with three seasonal human H1N1 strains (A/Brisbane/59/2007, A/Pennsylvania/08/2008 and A/Oklahoma/447/08), one human pandemic H1N1 strain (A/California/04/2009), and one H3N2 virus (A/Oklahoma/483/08) for comparison. Consistent with our previous
findings, A/Brisbane isolate showed the broadest binding specificity and preferred glycans with terminal α,2,6-linked sialic acid (Fig. 5A). Interestingly, the glycans recognized by A/Brisbane/59/2007 (H1N1) were the same glycans that bound SNA, which is specific for the determinant, Neu5Acα₂-6Galβ₁-4GlcNAc. The binding profile of A/Oklahoma/447/08 (H1N1) virus is similar to the A/Brisbane virus, but displayed a much higher signal to noise ratio (Fig. 5B). When compared with binding data from lectins and antibodies, A/Oklahoma/447/08 (H1N1) displayed a clear preference for glycans with α,2,6 sialic acid, binding strongly to 2,6-DS-NA₂, LSTc, and all of the glycans bound by SNA. Additionally, like A/Brisbane/59/2007, several glycans were recognized by the anti-SLe²/LSTA antibody, suggesting certain specificity toward α,3 sialic acid-containing glycans. A/Pennsylvania/08/2008, which was shown to preferentially bind glycans having terminal α,2,6 sialic acid when assayed on the CFG defined glycan array, differed from the other H1N1 strains, A/Oklahoma/447/08 and A/Brisbane/59/2007 in that it did not bind some of the HMGs (such as glycan H-12, H-15, H-16, H-53, and H-55) that possess the Neu5Acα₂,6Galβ₁-4 motif (Fig. 5C). This result suggests that the virus binding does not solely rely on the sialic acid linkage. In the case of A/California/04/2009, a human pandemic H1N1 isolate, the binding pattern overlaps with SNA (Fig. 5D), which confirms the results from CFG microarray that A/California/04/2009 has a restricted binding preference for α,2,6 sialic acid linked type-2 glycans. Further, as a comparison to the H1N1 A/Oklahoma/447/08 virus, we tested the A/Oklahoma/483/08 H3N2 virus isolate (Fig. 5E) and obtained a more restricted binding pattern where all the bound glycans contain Neu5Acα₂,6Galβ₁-4GlcNAc structure and were recognized by the H1N1 A/Oklahoma/447/08.

Finally, we evaluated the binding properties of three human parainfluenza viruses on the Human Milk SGM. It is known that these viruses require the presence of Neu5Acα₂-3Galβ₁-4GlcNAc motif (66). However, our MAL-I binding data showed that this structure is not a component of the HMG. Thus, as we would predict, no binding was observed with the type 1 and 3 (hPIV1 and hPIV3) viruses. By contrast, unlike type 1 and 3, the type 2 parainfluenza virus (hPIV2) displayed a very strict preference for α,2,6 sialic acid-containing glycans (Fig. 5F), with receptor specificity similar to the H1N1 strains. Overall, the virus binding experiments demonstrated that many of the HMGs might function as decoys for cell-bound receptors and that the elements in the HMGs might be found on cell-bound receptors. Glycans that bound viruses were selected for more detailed characterization as described below.

Structural characterization of selected HMGs – Interrogation of the partially characterized HMG-derived SGM with antibodies against biological markers and viruses demonstrated the potential to identify the receptors of GBPs, including lectin, antibodies, and GBPs in pathogens. Although the existing lectin/antibody binding data already provided some common features of the receptors, detailed structural analysis is necessary in order to relate the specific structures with biological functions. We selected relevant glycans from the Human Milk TGL and attempted the use of tandem mass spectrometry and/or serial enzymatic digestion (67) to decipher these structures and found that while MALDI analysis of the glycan derivatives generated excellent data and in some cases good secondary fragmentation data, more sophisticated analyses of permethylated glycan-AEAB derivatives were difficult to interpret due to the complexity of the spectra generated from partial methylation of the primary and secondary amines introduced by the AEAB. These complexities limited the detailed structural analysis of selected glycans from the TGL by MS/MS.

We realized, however, that our lectin-based analysis of the Human Milk SGM provided significant structural information in a rather high throughput format since all 127 glycans could be analyzed in a single assay. In addition we reasoned that digestion of the glycans on the microarray would be an efficient approach to do in situ structural analysis by combining specific exoglycosidase digestion with defined lectin
binding. To demonstrate this, we selected a total of 22 functionally identified glycans, and their HPLC profiles and MALDI-TOF spectra are shown in Figure S3. These structures include 7 glycans bound by anti-TRA-1 antibodies (H-71, H-99, H-108, H-111, H-125, H-126, and H-127), 6 glycans bound by MVM (H-28, H-30, H-31, H-36, H-43, and H-60) and 9 glycans bound by influenza viruses (H-1, H-5, H-6, H-12, H-14, H-15, H-16, H-17, and H-56). These glycans were printed as a separate array designated “HMG-subarray” on NHS-derivatized slides along with 18 structurally defined glycan standards. The 18 control glycans represent some typical structural motifs found in human milk, such as type-1, type-2 glycans and Lewis blood group glycans (structures shown in Fig. 6). The results from these glycans were used to monitor the behaviors of reagents and to direct structure predictions.

To obtain on-array sequence information, the non-reducing terminal structures of the selected glycans were first determined by screening the HMG-subarray with a variety of defined lectins and antibodies (listed in Table S3A) whose specificities were defined by analysis on v5.0 of the CFG defined glycan array. We reasoned that collection of this data, along with the predicted compositional data from mass spectrometry, could be combined as a collection of metadata and would provide information about the specific structures of glycans that mass spectrometry alone might not easily resolve. We have termed this approach Metadata-Assisted Glycan Sequencing (MAGS). To this end, positive/negative binding by lectins or antibodies to each glycan indicates the presence/absence of the corresponding moiety that each lectin or antibody recognizes, e.g. ECL for Galβ1-4GlcNAc, AAL for fucose and anti-Lea antibody for Lea epitope. The binding data from multiple GBPs were analyzed in detail to assign the structures. For example, if SNA, RCA-I and anti-type-1 chain antibody, but not ECL showed binding toward a glycan, it would suggest that this glycan might possess a terminal Neu5Acα2-6Galβ1-4GlcNAc determinant (SNA and RCA-I positive; ECL negative) together with a terminal Galβ1-3GlcNAc determinant (ECL negative and anti-Type-1 antibody positive). These data do not distinguish between an asymmetric biantennary glycan and the presence of two glycans; however, other metadata can be associated with the individual printed glycans such as a MALDI analysis to determine the number of molecular ions and composition and the shape of individual peak(s) during HPLC to evaluate glycan homogeneity. The next series of experiments involved the use of a group of specific exoglycosidases to treat the glycans directly on the microarray followed by interrogating with defined lectins or antibodies. The gain and loss of binding can provide composition and/or linkage information for the terminal and penultimate sugar residues. To accomplish this we optimized the reaction condition of five exoglycosidases (Table S3B) for on-array digestion, including the non-specific neuraminidase from *Arthrobacter ureafaciens*, the recombinant α2,3-neuraminidase from *Salmonella typhimurium*, the jack bean β1-4/6 galactosidase, the recombinant β1,3-galactosidase from *Xanthomonas manihotis* and the recombinant β1,4-galactosidase from *Bacteroides fragilis*. In general, we found that longer incubation times and high enzyme concentrations were necessary to achieve effective digestion when compared with the reactions in solution. The binding results after enzymatic treatment can be divided into two categories, loss and gain. The loss of signal after digestion confirms the prediction from the positive signals before digestion, as in the case of neuraminidase treatment, where the loss of SNA or anti-SLea antibody binding confirms the presence of a specific sialic acid linkage. Similarly, β-galactosidase digestion confirms a type-1 or type-2 chain structure. For type-1 chain, the binding of anti-type-1 chain antibody is specifically diminished by β1,3-galactosidase treatment, while β1,4-galactosidase has no effect. On the contrary, the binding of ECL to a type-2 chain is lost only after β1,4-galactosidase digestion. Furthermore, the type-1 and type-2 structures can also be distinguished by the gain of GSL-II binding, after β1-3/4-galactosidase digestion removes a terminal Gal revealing a terminal GlcNAc. Beyond the single enzymatic treatment, we also conducted sequential digestion involving desialylation first with neuraminidase or α2,3-
specific neuraminidase followed by specific β-galactosidase treatment. This set of experiment is particularly useful for sialylated glycans and HMGs that are comprised of many isomers of type 1 and type 2 linear and branched lactosamines and poly-N-acetyllactosamines. Finally, all of the collected metadata including molecular ions, fragmentation MS or MS/MS data, and the behavior on ion exchange chromatography and HPLC are combined to provide predictions of structures that were not possible with MALDI-TOF and MALDI-TOF/TOF analyses alone.

Using this MAGS approach, we were able to predict the structural moieties and most of the linkages for 20 of the 22 HMGs as shown in Figure 6. The structures of H-06 and H-71 were not proposed since these glycans appeared to be mixtures. The binding data is compiled in Table S4 and the lectin binding and specific antibody binding data are summarized in Tables 1 and 2, respectively. The detailed description of the logic used to predict the structure of each glycan is provided in the supplemental material. Six pairs of glycans were found to be the same structure (H1/H5, H14/H16, H15/H17 and H28/H30) or have the same general structures (H43/H60 and H111/H126) based on HPLC profiles, MALDI analysis, and binding data. This is due to the overlap of glycans in the fractions obtained during the multi-dimensional chromatography. In addition, several samples were contaminated with minor impurities. Nevertheless, the correlation of proposed structures with the function defined by antibody and virus binding revealed interesting findings. The five influenza virus receptors, including monosialylated H-01/05, H12, H14/16, H15/17 and disialylated H-56 all contain the Neu5Acα2,6Galβ1-4Glc/GlcNAc moiety as indicated by SNA, RCA-1, ECL, neuraminidase and β1-4 galactosidase data (Table 1). Except for the sialyl lactose (H01/05), the other four structures are biantennary glycans with one type-2 chain branch. It appears that the other branch can be diverse structures, as we observed the presence of Leα, Leα, type-2 chain, and sialylated type-1 chain. The on-array structural analysis also revealed common features for the MVM receptors. Mostly relying on antibodies (anti-SLeα/LSTα and anti-type-1) and exoglycosidases (specific and unspecific neuraminidase) data (Table 2), we proposed that the four disialyl structures all carry an α2,3-sialylated type-1 chain with an additional sialic acid attached to the GlcNAc in α2,6 linkage (Neu5Acα2,3Galβ-3(Neu5Acα2,6)GlcNAc). It is possible that the disialyl LNT motif is one of the recognition determinants for MVM. Similar to influenza virus, modifications such as fucosylation and branching on this motif did not block the virus recognition. Further, our results, together with the CFG data, indicated that the recognition of MVMs is beyond the sialic acid since the viruses did not bind to all the multisialylated glycans. In the case of anti-TRA-1 antibodies, the sialylated binder H-71 was found to have relatively low purity and thus its structure was not elucidated. The other 6 binders (H-99, H-108, H-111, H-125, H-126, and H-127) are neutral complexed glycans, especially for the latter four, which are multi-branched structures. Although we did not obtain all of the linkage information for these large glycans, we found that all the receptors contain the common motif: type-1 lactosamine epitope. While H-99 is a relatively simple lactosamine glycan, similar to the structure identified from CFG array, the others contain an additional type-2, Leα or Leβ branch and it seems that these extra branches do not prevent the binding of the antibodies.

DISCUSSION

Human milk glycans have been studied for many decades and many components have been purified and structurally characterized. It is recognized that HMGs play multiple biological roles and probably each function is directly correlated with specific structures of individual glycans (6,8). However, there has been no efficient method for functional glycomic analyses of HMGs in terms of their recognition and interactions with other molecules. Using a shotgun glycomics approach, we isolated a human milk free glycan glycome and generated a corresponding glycan microarray that allowed us to investigate the function of individual glycans in a high-throughput format. The binding patterns of defined lectins and antibodies to the Human Milk SGM revealed the
The structural diversity of the isolated glycome. The Human Milk TGL, which represents a permanent repository for HMG analyses, is enriched in sialylated glycans, type-1 and type-2 glycans, and Le$^a$-containing glycans, but lacks α1,2-linked fucosylated glycans, such as H type-1 and Le$^b$ structures. These data indicate that the milk sample used for preparation of the tagged glycan library was from a Lewis positive individual and confirm that the individual was a secretor negative donor. We chose to use milk from a secretor negative donor in this first study, since it would limit the diversity of glycans available for interrogation. In further studies we are preparing Human Milk SGMs from donors of other genotypes. It is clear from our studies that analyses of glycans of individual milk samples on microarrays using specific blood group antibodies may represent an alternative strategy to the existing mass-based (68) or HPAEC-based methods (4) for screening the genotype of milk donors. The Human Milk SGM described here provides a source of purified glycans that are conveniently available as an archived resource for interrogation with antibodies, pathogenic microorganisms and potentially live cells for evaluating functions of individual structures. In this approach, once a biologically relevant structure is identified, many of its structural features can be described by collecting metadata in a database for the entire shotgun array. If more detailed analysis; i.e., actual sequence and definition of complete structure are required, the preliminary structural features will be invaluable for analysts using other physical techniques such as Mass spectrometry and NMR. For example, monoclonal antibodies, TRA-1-60 and TRA-1-81 are widely used to identify the biomarker for human pluripotent stem cells. These antibodies were generated using cells as antigens, and they were originally thought to be directed against keratan-sulfate proteoglycan (69). In recent studies using the defined glycan array from the CFG (version 4.2), Natunen et al. identified type-1 lactosamine as an epitope for TRA-1-60 and TRA-1-81 antibodies (63). Here, we extended those observations to additional structures that were identified first as glycans recognized by these antibodies followed by determining their structure using metadata collected on these specific glycans on the shotgun array. Our results revealed the presence of ligands for the anti-TRA-1 antibodies in the human milk glycome and showed that the two stem cell marker antibodies recognize complicated large glycans that possess the reported minimum recognition moiety, type-1 lactosamine. Although there is no obvious correlation between the stem cell marker antibodies and the function of HMGs, we demonstrated that Human Milk SGM is a glycan library useful for investigating the binding specificity of proteins due to the diversity and complexity of HMGs, and especially due to its enrichment of blood group related epitopes, which are the receptor of various proteins and microorganisms. Another recent example is the newly generated anti–stage-specific embryonic antigen (SSEA)-5 monoclonal antibody, which was found to only recognize H type-1 glycans (70) that are characteristic of HMGs.

Inhibition of pathogen binding is considered the essential immune protection provided to the infant by human milk. It is believed that certain HMGs act as the receptor analogues or decoy receptors to bind to pathogens and thus prevent their attachment to cell targets (3). The established Human Milk SGM provides a unique platform to study the interaction between pathogens and milk oligosaccharides. To demonstrate the ability of the Human Milk SGM for evaluating the binding specificity of pathogens, we tested MVM and influenza virus, both known to recognize sialic acids, which are enriched in the Human Milk SGM. It was interesting that our study revealed that MVMs specifically bound to α2,3-linked multi-sialyl glycans, which are a category of novel receptors that are not present on the CFG microarray. Although there is no connection between the HMG function and MVMs, these results could provide additional insight into the receptor preferences of MVMs, which were previously found to recognize α2,8-linked mutisialylated glycans and sialyl Le$^x$ moieties on the CFG microarray (41). Combining the study from both glycan arrays, it is clear that MVMs have a unique preference for certain multi-sialylated
oligosaccharides, like disialyl LNT glycans we identified in this paper.

To expand our study to human influenza viruses beyond the synthetic CFG microarray (37) and the sialylated glycan microarray (71), we interrogated the Human Milk SGM with several influenza strains. Similar to previous findings, the viruses have a strong preference toward α2,6-sialylated glycans, but not all the α2,6-sialylated glycans were recognized by the viruses, which once again suggested that the recognition of influenza viruses is not limited only to the terminal glycans. One common feature for the receptors is that they all contain an α2,6-linked sialic acid to a type-2 chain. Overall, our virus experiments demonstrated that the Human Milk SGM is useful for studying pathogen binding properties, and presumably possesses receptors for pathogens associated with infant diseases, which will be the focus of future studies.

A major tool used in the current work was the combination of metadata from mass spectrometry, lectin and antibody binding, and on-array enzymatic modification degradation of glycans in the technique we termed MAGS. This approach can complement the use of mass spectrometry and other approaches and provides key information important in predicting glycan structure. Development of MAGS partly arose from our needs to identify the recognition motifs and even underlying glycan structures required for better understanding the interactions of glycans with proteins and microorganisms. Compared to the low cost and high efficiency of nucleic acid and protein sequencing analyses, glycan sequence analysis is very difficult and often requires highly sophisticated instrumentation and skilled interpretation. Although progress in mass spectrometry has significantly advanced the structural analysis of glycans, there is a need to develop more high-throughput and fast glycan sequencing methods that only require small amounts of samples. In this work here, we initiated the on-array glycan sequencing method of MAGS and chose Human Milk SGM as the working model.

Using lectins and antibodies whose binding specificities are known, we collected binding data from original glycans and enzyme-treated glycans. Various sugar moieties and linkage information were deduced from analyzing all of the data. Extremely minute amounts of HMGs were used and many individual glycans were analyzed simultaneously on a single microarray. These studies demonstrated the successful use of the on-array glycan sequencing method and also showed that this technique is currently limited only by the availability of lectins and antibodies with defined binding specificities, as well as specific exoglycosidases. For example, there is no reagent that can differentiate the branch-point (β1,3/6) to the core lactose/lactosamine, although the current literature indicates that the lactosamine unit at this β1-6 branch-point is a type-2 structure (21). Except for neuraminidase, the enzymes used showed relative low activity on the slides comparing to the activity in solution and thus required longer incubation time and relatively large amounts of enzymes (Table S3). Clearly, the analyses described here could be extended by the use of specific α-L-fucosidases, but these studies were limited by their availability and cost from commercial sources. Nevertheless, using the on-array sequencing analysis, we were able to successfully identify many of the structural features important for antibody binding and the receptors for several viruses in a high throughput format.

Collecting milk glycan structures and having methods to rapidly identify milk glycans is of course useful, and can generate interesting data on the metabolic changes of specific milk glycans, but generally do not provide direct information on glycan function. Having glycans immobilized on a microarray for interrogation by biologically relevant GBPs or microorganisms and retrievable from a TGL provides a true approach to functional glycomic analysis to accelerate discovery of HMG function. As an added feature, collecting metadata on each glycan associated with a SGM as the array is interrogated with both defined GBPs and GBPs of interest to reveal glycan function, these arrays of immobilized glycans can be structurally characterized in a high
throughput format. As more defined GBPs and highly purified exoglycosidases become available, this MAGS approach could be miniaturized and automated as a high throughput analytical tool capable of sequencing thousands of glycans simultaneously.

References


Comparison of the receptor binding properties of contemporary swine isolates and early human pandemic H1N1 isolates (Novel 2009 H1N1). *Virology* 413, 169-182


**Abbreviations** — The abbreviations used are: HMGs, human milk glycans; GBP, glycan binding protein; AEAB, 2-amino-N-(2-aminoethyl)-benzamide; Hex, hexose; HexNAc, N-acetylhexosamine; Fuc, fucose; Neu5Ac/SA, sialic acid; NHS, N-hydroxysuccinimide; Con A, concanavalin A; VVL, *Vicia villosa* lectin; GSL-II, *Griffonia simplicifolia* lectin II; MAL-I, *Maackia amurensis* Lectin I; AAL, *Aleuria aurantia* lectin; SNA, *Sambucus nigra* agglutinin; LTL, *Lotus tetragonolobus* lectin; UEA-I, *Ulex europaeus* agglutinin I; RCA-I, *Ricinus communis* agglutinin I; ECL, *Erythrina cristagalli* Lectin; RFU, relative
fluorescence units; Le\textsuperscript{a}, Lewis\textsuperscript{a}; Le\textsuperscript{b}, Lewis\textsuperscript{b}; Le\textsuperscript{c}, Lewis\textsuperscript{c}; Le\textsuperscript{d}, Lewis\textsuperscript{d}; CFG, Consortium for Functional Glycomics; MVM, minute virus of mice; hPIV, human parainfluenza virus. The structures corresponding to abbreviated glycans are defined in Fig. 6 and Table S4.

**Acknowledgments** — We thank Shelly Gulati (University of Oklahoma Health Sciences Center) for her help on influenza virus preparations and Dr. Jamie Heimburg-Molinaro (Emory University School of Medicine) for manuscript editing and review.

**Footnotes** — *This work was supported by National Institute of Health Grants GM62116 (Consortium for Functional Glycomics) and RO1GM085448 (DFS), National Science Foundation grant MCB 0718948 (MA-M), and USDHHS contract HHSN266200700006C (NIAID Centers of Excellence for Influenza Research and Surveillance).
FIGURE LEGENDS

Figure 1. Schematic for generation of Human Milk Glycan (HMG) Shotgun Glycan Microarray (SGM). Human milk oligosaccharides were extracted, fractionated, AEAB conjugated (labeled with a tag) and separated. The purified fractions were quantified and printed to create a Human Milk Glycan SGM available for studies with GBPs and microorganisms.

Figure 2. Plant lectins binding to Human Milk SGM. The Human Milk SGM microarray was characterized with biotinylated lectins: (A) AAL (0.1 µg/ml), (B) UEA-I (10 µg/ml), (C) LTL (10 µg/ml), (D) SNA (5 µg/ml), (E) RCA-I (10 µg/ml) and (F) ECL (10 µg/ml). A total of 140 glycans were printed on the microarray. Glycans 1-73 are sialylated glycans (tan); 74-127 are neutral glycans (violet); 128-140 are controls of structurally defined glycans (light green). The structures in symbols indicate the binding specificity of each lectin identified by defined glycan microarray (CFG v5.0). Considering the variation of binding affinity, the histogram shows the data at the concentration that yielded the best signal/background ratio instead of the data at the same concentration.

Figure 3. Antibodies binding to Human Milk SGM. The Human Milk SGM was interrogated with antibodies: (A) anti-Le^a antibody (10 µg/ml), (B) anti-SLe^a/LSTa antibody (10 µg/ml), (C) anti-blood group H1 antibody (1:10 dilution) and (D) anti-CD15 antibody (10 µg/ml). The microarray was also used to test the binding specificity of (E) anti-TRA-1-60 antibody (50 µg/ml) and (F) anti-TRA-1-81 antibody (50 µg/ml). Glycans 1-73 are sialylated glycans (tan); 74-127 are neutral glycans (violet); 128-140 are controls of structurally defined glycans (light green).

Figure 4. MVM viruses binding to Human Milk SGM. The binding preferences of several strains of minute virus of mice (MVM): (A) MVMP-WT (prototype strain, empty capsid), (B) MVMP-VLP (prototype strain, virus-like particles), (C) MVMI-ggA (capsid protein mutant of immunosuppressive strain, empty capsid), and (D) MVMI-agD (Non-structural protein mutant of immunosuppressive strain, empty capsid) were evaluated on the Human Milk SGM. Each virus was tested at 200 µg/ml and detected by anti-MVM capsid antibody. Panel (E) shows that there was no background binding from the anti-MVM capsid antibody to HMG microarray. Glycans 1-73 are sialylated glycans (tan); 74-127 are neutral glycans (violet); 128-140 are controls of structurally defined glycans (light green).

Figure 5. Influenza viruses binding to Human Milk SGM. The binding preferences of various influenza virus isolates: (A) A/Brisbane/59/2007 H1N1, (B) A/Oklahoma/447/2008 H1N1, (C) A/Pennsylvania/08/2008, (D) A/California/04/2009 H1N1, (E) A/Oklahoma/483/2008 H3N2, and (F) hPIV2 were evaluated with HMG microarray. Glycans 1-73 are sialylated glycans (tan); 74-127 are neutral glycans (violet); 128-140 are controls of structurally defined glycans (light green).

Figure 6. The structures of 17 defined glycans used for controls and the predicted structures of 20 selected human milk oligosaccharides. The control glycans (left panel) are listed by common names and the HMGs (panels A, B, C) are listed by fraction names that are the same as the glycan ID on the microarray. Glycan H01-H17 and H-56 were ligands of influenza viruses; glycans H-28-H-36 were ligands of MVMs; glycans H-99-H-127 were ligands of anti-TRA-1 antibodies. 18 control glycans were printed, however glycans 11 and 13 were both LNFP.I.
TABLE LEGENDS

Table 1. Heat map of lectin binding data for structural analysis. A heat map summary of the binding data from analyses of the HMG-subarray with different plant lectins before and after specific exoglycosidase digestion(s). A total of 40 glycans were printed on the HMG-subarray, including 18 control glycans and 22 selected HMGs. The color-coded numbers highlight the binding intensity (the darker the color, the higher the binding intensity). The color scale for the Excel spreadsheet was set using 0 as minimum and 60,000 RFU as maximum. The untreated HMG-subarrays were tested with eight lectins. The HMG-subarrays treated with non-specific neuraminidase and α2-3-specific neuraminidase were tested with SNA and ECL. The β1-3 galactosidase (β1-3 galase), or β1-4 galactosidase (β1-4 galase, from either NEB or Prozyme) treated slides were tested with GSL-II. The lectin concentrations used were 10 µg/ml with the exception of AAL, which was 1 µg/ml.

Table 2. Heat map of antibody binding data for structural analysis. A heat map summary of the binding data from analyses of the HMG-subarray with different defined antibodies before and after specific exoglycosidase digestion(s). A total of 40 glycans were printed on the HMG-subarray, including 18 control glycans and 22 selected HMGs. The color-coded numbers highlight the binding intensity (the darker the color, the higher the binding intensity). The color scale for the Excel spreadsheet was set using 0 as minimum and 1,500 RFU as maximum for anti-Leα, anti-type1 and anti-CD15 antibodies; and 0 as minimum and 10,000 RFU as maximum for anti-SLeα and Leα antibodies. The untreated HMG-subarrays were tested with 5 antibodies. The HMG-subarrays treated with non-specific neuraminidase were interrogated with anti-Leα, anti-type1, anti-SLeα, anti-CD15, and the HMG-subarrays treated with α2-3-specific neuraminidase were interrogated with anti-Leα, anti-type1, and anti-SLeα.
Yu, et al. Figure 1
Yu, et al. Figure 2
Yu, et al. Figure 3
Control glycans

A. Influenza receptors

B. MVM receptors

C. Epitopes for anti-TRA antibody

Yu, et al Figure 6
Table 1. Heat map of lectin-binding data for structural analysis

<table>
<thead>
<tr>
<th>Print ID</th>
<th>Glycan Name</th>
<th>Composition</th>
<th>Neuraminidase</th>
<th>2,3-Neuraminidase</th>
<th>β1,3-galactose</th>
<th>β1,4-galactose</th>
<th>Neuraminidase</th>
<th>2,3-Neuraminidase</th>
<th>β1,3-galactose</th>
<th>β1,4-galactose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NA2</td>
<td>122</td>
<td>61111</td>
<td>54859</td>
<td>59</td>
<td>17</td>
<td>60496</td>
<td>61</td>
<td>28</td>
<td>14856</td>
</tr>
<tr>
<td>2</td>
<td>2,3-DesNA2</td>
<td>67</td>
<td>143</td>
<td>15</td>
<td>62</td>
<td>11</td>
<td>61270</td>
<td>53</td>
<td>413</td>
<td>28736</td>
</tr>
<tr>
<td>3</td>
<td>2,6-DesNA2</td>
<td>56355</td>
<td>60330</td>
<td>24</td>
<td>57</td>
<td>10</td>
<td>59469</td>
<td>53</td>
<td>29</td>
<td>17644</td>
</tr>
<tr>
<td>4</td>
<td>LNT</td>
<td>78</td>
<td>7246</td>
<td>94</td>
<td>112</td>
<td>22</td>
<td>64</td>
<td>34</td>
<td>20</td>
<td>140</td>
</tr>
<tr>
<td>5</td>
<td>AgalLNT</td>
<td>69</td>
<td>201</td>
<td>40</td>
<td>57</td>
<td>21851</td>
<td>96</td>
<td>49</td>
<td>49</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>LNAaT</td>
<td>71</td>
<td>58508</td>
<td>28867</td>
<td>68</td>
<td>22</td>
<td>-27</td>
<td>52</td>
<td>58</td>
<td>11011</td>
</tr>
<tr>
<td>7</td>
<td>LD, Lea</td>
<td>47</td>
<td>131</td>
<td>8</td>
<td>59766</td>
<td>33</td>
<td>154</td>
<td>57357</td>
<td>91</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>LSTA</td>
<td>82</td>
<td>110</td>
<td>29</td>
<td>70</td>
<td>26</td>
<td>-18</td>
<td>59</td>
<td>25</td>
<td>133</td>
</tr>
<tr>
<td>9</td>
<td>LSB</td>
<td>78</td>
<td>1283</td>
<td>24</td>
<td>78</td>
<td>24</td>
<td>-23</td>
<td>62</td>
<td>46</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>LSTc</td>
<td>53314</td>
<td>26175</td>
<td>34</td>
<td>76</td>
<td>23</td>
<td>-4</td>
<td>64</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>LNFPI</td>
<td>74</td>
<td>156</td>
<td>25</td>
<td>55253</td>
<td>17</td>
<td>-13</td>
<td>85</td>
<td>31</td>
<td>24</td>
</tr>
<tr>
<td>12</td>
<td>LNFPi, Lea</td>
<td>69</td>
<td>151</td>
<td>15</td>
<td>60099</td>
<td>26</td>
<td>28</td>
<td>1730</td>
<td>38</td>
<td>16</td>
</tr>
<tr>
<td>13</td>
<td>LNFPi</td>
<td>48</td>
<td>183</td>
<td>40</td>
<td>56267</td>
<td>39</td>
<td>38</td>
<td>82</td>
<td>60</td>
<td>24</td>
</tr>
<tr>
<td>14</td>
<td>LNFPIi, Lea</td>
<td>84</td>
<td>128</td>
<td>33</td>
<td>60075</td>
<td>4937</td>
<td>-6</td>
<td>61</td>
<td>43</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>Sambu, Lea</td>
<td>87</td>
<td>148</td>
<td>16</td>
<td>59462</td>
<td>15</td>
<td>41</td>
<td>84</td>
<td>44</td>
<td>23</td>
</tr>
<tr>
<td>16</td>
<td>Sambu, Lea</td>
<td>80</td>
<td>121</td>
<td>16</td>
<td>59784</td>
<td>8</td>
<td>12</td>
<td>57</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>DSNT</td>
<td>90</td>
<td>178</td>
<td>58</td>
<td>63</td>
<td>17</td>
<td>-16</td>
<td>53</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>18</td>
<td>Lea, Lea</td>
<td>58</td>
<td>127</td>
<td>15</td>
<td>57946</td>
<td>92</td>
<td>93</td>
<td>1852</td>
<td>44</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>H-01</td>
<td>2759</td>
<td>3749</td>
<td>66</td>
<td>104</td>
<td>58</td>
<td>-12</td>
<td>86</td>
<td>40</td>
<td>24</td>
</tr>
<tr>
<td>20</td>
<td>H-02</td>
<td>1900</td>
<td>6220</td>
<td>65</td>
<td>6240</td>
<td>24</td>
<td>-19</td>
<td>70</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td>21</td>
<td>H-06</td>
<td>7472</td>
<td>4268</td>
<td>73</td>
<td>277</td>
<td>87</td>
<td>-11</td>
<td>96</td>
<td>90</td>
<td>7</td>
</tr>
<tr>
<td>22</td>
<td>H-12</td>
<td>54882</td>
<td>7859</td>
<td>31</td>
<td>62142</td>
<td>80</td>
<td>82</td>
<td>66</td>
<td>58</td>
<td>20</td>
</tr>
<tr>
<td>23</td>
<td>H-14</td>
<td>50209</td>
<td>59216</td>
<td>10497</td>
<td>246</td>
<td>30</td>
<td>266</td>
<td>74</td>
<td>80</td>
<td>18</td>
</tr>
<tr>
<td>24</td>
<td>H-15</td>
<td>55202</td>
<td>32504</td>
<td>24</td>
<td>20946</td>
<td>26</td>
<td>64</td>
<td>64</td>
<td>68</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>H-16</td>
<td>51341</td>
<td>60496</td>
<td>16025</td>
<td>81</td>
<td>11</td>
<td>13</td>
<td>62</td>
<td>30</td>
<td>21</td>
</tr>
<tr>
<td>26</td>
<td>H-17</td>
<td>59186</td>
<td>50444</td>
<td>42</td>
<td>61695</td>
<td>41</td>
<td>22</td>
<td>73</td>
<td>36</td>
<td>24</td>
</tr>
<tr>
<td>27</td>
<td>H-28</td>
<td>53315</td>
<td>26175</td>
<td>34</td>
<td>76</td>
<td>23</td>
<td>-4</td>
<td>64</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>28</td>
<td>H-30</td>
<td>70</td>
<td>89</td>
<td>26</td>
<td>24800</td>
<td>20</td>
<td>-10</td>
<td>55</td>
<td>35</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>H-31</td>
<td>93</td>
<td>110</td>
<td>3417</td>
<td>21</td>
<td>-33</td>
<td>71</td>
<td>10</td>
<td>15</td>
<td>81</td>
</tr>
<tr>
<td>30</td>
<td>H-36</td>
<td>6419</td>
<td>26175</td>
<td>3332</td>
<td>165</td>
<td>47</td>
<td>19</td>
<td>84</td>
<td>79</td>
<td>6</td>
</tr>
<tr>
<td>31</td>
<td>H-56</td>
<td>94</td>
<td>112</td>
<td>68</td>
<td>57231</td>
<td>40</td>
<td>26</td>
<td>67</td>
<td>44</td>
<td>23</td>
</tr>
<tr>
<td>32</td>
<td>H-56</td>
<td>61399</td>
<td>60861</td>
<td>352</td>
<td>18668</td>
<td>60</td>
<td>3</td>
<td>99</td>
<td>80</td>
<td>26</td>
</tr>
</tbody>
</table>

Yu, et al., Table 1
### Table 2. Heat map of antibody-binding data for structural analysis

<table>
<thead>
<tr>
<th>Print ID</th>
<th>Glycan Name</th>
<th>Composition</th>
<th>Neuraminidase</th>
<th>2.3-neuraminidase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Le^a</td>
<td>anti-type1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Nle</td>
<td>anti-SLe^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-CD15</td>
<td>anti-Le^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-type1</td>
<td>anti-Nle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-SLe^a</td>
<td>anti-CD15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Le^a</td>
<td>anti-type1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Nle</td>
<td>anti-SLe^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-CD15</td>
<td>anti-Le^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-type1</td>
<td>anti-Nle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-SLe^a</td>
<td>anti-CD15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Print ID</th>
<th>Glycan Name</th>
<th>Composition</th>
<th>Neuraminidase</th>
<th>2.3-neuraminidase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Le^a</td>
<td>anti-type1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Nle</td>
<td>anti-SLe^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-CD15</td>
<td>anti-Le^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-type1</td>
<td>anti-Nle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-SLe^a</td>
<td>anti-CD15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Le^a</td>
<td>anti-type1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-Nle</td>
<td>anti-SLe^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-CD15</td>
<td>anti-Le^a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-type1</td>
<td>anti-Nle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>anti-SLe^a</td>
<td>anti-CD15</td>
</tr>
</tbody>
</table>


Yu et al., Table 2
Functional glycomic analysis of human milk glycans reveals presence of virus receptors and embryonic stem cell biomarkers
Ying Yu, Shreya Mishra, Xuezhe Song, Yi Lasanajak, Konrad C. Bradley, Mary M. Tappert, Gillian M. Air, David A. Steinhauer, Sujata Halder, Susan Cotmore, Peter Tattersall, Mavis Agbandje-McKenna, Richard D. Cummings and David F. Smith

J. Biol. Chem. published online October 31, 2012

Access the most updated version of this article at doi: 10.1074/jbc.M112.425819

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2012/10/31/M112.425819.DC1