Novel lamprey antibody recognizes terminal sulfated galactose epitopes on mammalian glycoproteins

Tanya R McKitrick, Harvard Medical School
Steffen M Bernard, The Scripps Research Institute, La Jolla
Alexander J Noll, Harvard Medical School
Bernard C Collins, Scripps Research Institute
Christopher K Goth, Harvard Medical School
Alyssa M McQuillan, Harvard Medical School
Jamie Heimburg-Molinaro, Harvard Medical School
Brantley R Herrin, Emory University
Ian A Wilson, The Scripps Research Institute, La Jolla
Max Cooper, Emory University

Only first 10 authors above; see publication for full author list.

Journal Title: COMMUNICATIONS BIOLOGY
Volume: Volume 4, Number 1
Publisher: NATURE RESEARCH | 2021-06-03, Pages 674-674
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1038/s42003-021-02199-7
Permanent URL: https://pid.emory.edu/ark:/25593/vxnfg

Final published version: http://dx.doi.org/10.1038/s42003-021-02199-7

Copyright information:
© The Author(s) 2021

This is an Open Access work distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/rdf).

Accessed December 30, 2023 8:33 PM EST
Novel lamprey antibody recognizes terminal sulfated galactose epitopes on mammalian glycoproteins

Tanya R. McKitrick, Steffen M. Bernard, Alexander J. Noll, Bernard C. Collins, Christoffer K. Goth, Alyssa M. McQuillan, Jamie Heimburg-Molinaro, Brantley R. Herrin, Ian A. Wilson, Max D. Cooper & Richard D. Cummings

The terminal galactose residues of N- and O-glycans in animal glycoproteins are often sialylated and/or fucosylated, but sulfation, such as 3-O-sulfated galactose (3-O-SGal), represents an additional, but poorly understood modification. To this end, we have developed a novel sea lamprey variable lymphocyte receptor (VLR) termed O6 to explore 3-O-SGal expression. O6 was engineered as a recombinant murine IgG chimera and its specificity and affinity to the 3-O-SGal epitope was defined using a variety of approaches, including glycan and glycoprotein microarray analyses, isothermal calorimetry, ligand-bound crystal structure, FACS, and immunohistochemistry of human tissue macroarrays. 3-O-SGal is expressed on N-glycans of many plasma and tissue glycoproteins, but recognition by O6 is often masked by sialic acid and thus exposed by treatment with neuraminidase. O6 recognizes many human tissues, consistent with expression of the cognate sulfotransferases (GAL3ST-2 and GAL3ST-3). The availability of O6 for exploring 3-O-SGal expression could lead to new biomarkers for disease and aid in understanding the functional roles of terminal modifications of glycans and relationships between terminal sulfation, sialylation and fucosylation.
**G**lycosylation is an essential process that occurs in all known organisms and mediates multiple biological processes such as cell–cell interactions, cellular integrity, cellular signaling, and immune functions. Glycans on glycoproteins and glycolipids, as well as glycosaminoglycans, can be further modified by the addition of phosphates or sulfates to particular glycan residues, which can result in divergent functional outcomes. Terminal modifications of glycans are also important in their functions, and those, such as sialylation and fucosylation, are found on terminal galactose (Gal) or penultimate N-acetylgalactosamine (GlcNAc) residues. The addition of a negatively charged moiety, including sulfate or phosphate, to a glycan can have a substantial impact on the physiological landscape of the cell membrane, and can contribute significantly to protein–glycan interactions.

Sulfated glycans in particular play diverse roles, and the loss of these modifications can result in disease or altered functions.

A particularly enigmatic modification is 3-O-sulfation of galactose (3-O-Sgal) generated by members of the Gal3ST family, which use 3-phosphoadenosine-5’-phosphosulfate as the donor. 3-O-Sgal is found in the glycosphingolipid sulfatide, as generated by Gal3ST1. Such sulfation is essential for central nervous system maintenance, as shown by loss of myelin function and oligodendrocyte differentiation observed in Gal3ST1 knockout animals, as well as being required for germ cell functions.

There is much less understanding, however, about the role(s) of 3-O-Sgal in glycoproteins, where it has been reported that this modification is expressed in several glycoproteins, such as thyroglobulin, thyroid stimulating hormone, bovine peripheral myelin glycoprotein P0, and mouse brain neural cell adhesion molecule. The presence of 3-O-Sgal is regulated by Gal3ST-2 and Gal3ST-3, both of which can sulfate type-2 LacNAc (Galβ1-4GlcNAc) on N-glycans, whereas Gal3ST-4 prefers Galβ1-3GalNAc on O-glycans. Despite its importance as a terminal glycan modification, 3-O-Sgal has been difficult to study and analyze because it is relatively labile and analysis commonly require special procedural modifications.

There are no reagents for directly exploring 3-O-Sgal expression, thus its frequency and distribution in mammalian tissues and cells are poorly described. Interestingly, some approaches may miss the expression of 3-O-Sgal, due to cross-reactivity of some plant lectins that are thought to bind only sialic acid, but can also bind to 3-O-Sgal.

To better define the expression of 3-O-Sgal, we leveraged the unique adaptive immune system that has evolved in the sea lamprey, Petromyzon marinus. As antigen receptors, the lamprey utilize a family of highly diverse proteins called variable lym-

The VLRB germline sequence is incomplete, and is flanked by numerous LRR modules that are incorporated into the mature VLRB transcript via a specialized gene conversion-like mechanism that is capable of generating >1014 distinct receptors. The resulting VLRB protein is a 15–25 kDa, crescent-shaped protein with a continuous beta sheet where its concave surface forms the antigen binding site. Variation in antigen recognition is achieved by amino-acid sequence diversity located within the antigen binding site as well as the number of LRR segments that are incorporated into the mature protein.

Sea lampreys can induce robust VLRB responses to multiple glycan determinants upon immunization with a variety of glycan-bearing immunogens. In particular, analysis of immunized plasma on glycan microarrays revealed that lampreys can generate VLRBs specific for a variety of sulfated glycans, regardless of the immunogen. Herein, we report on the characterization of a lamprey VLRB that is specific for terminal 3-O-Sgal, and was isolated from a lamprey immunized with human type O erythrocytes. This VLRB, termed O6, is the first reagent to our knowledge that specifically recognizes the 3-O-Sgal determinant on glycoproteins with high sensitivity. The use of O6 to identify the presence of 3-O-Sgal on blood glycoproteins and in different cells and tissues, and its masking by sialylation as shown here, demonstrate that O6 will be useful for glycomic analyses and for characterizing expression profiles and roles of 3-O-Sgal in health and disease.

**Results**

**Isolation, identification, sequencing, and cloning of O6.** We identified O6 as an antibody that recognizes 3-O-Sgal using a yeast surface display (YSD) library generated from immunization of lampreys with type O human erythrocytes. Using methodology previously described, an amplified VLRB-specific cDNA library was transfected into the Saccharomyces EBY100 yeast strain with the pC7-ESO-BDNF expression vector. Using this expression platform, individual VLRB proteins are induced to the cell surface by a galactose promoter, where they are tethered to the endogenous cell surface Aga1p and Aga2p proteins. To identify antigen-specific VLRB clones, we incubated the induced VLRB YSD library with biotinylated intact erythrocytes and enriched by positive selection using streptavidin coupled magnetic beads (MACS). We further enriched this library by FACS of the VLRB-induced yeast bound to erythrocytes. This enriched yeast library was incubated on an extensive glycan microarray from the Consortium for Functional Glycomics (CFG), where the induced VLRBs on yeast can bind directly to the sugars printed on the microarray surface. After extensive washing, the bound clones were transferred to SD-CAA-agar plates and the VLRBs were sequenced. Yeast clones with unique VLRB sequences were individually screened by reexamination on the CFG array, and bound yeast clones were detected with an anti-Myc secondary antibody. The O6 VLRB was chosen because it exhibited unique specificity to 3-O-Sgal, as described below, and was then permanently transduced into HEK293 cells via a lentiviral construct to produce the recombinant VLRB mouse IgG2a (O6-mFc) fusion protein.

**Screening of O6 on glycan microarrays.** As previously mentioned, the O6 VLRB was discovered as a yeast clone bound to glycans containing 3-O-Sgal on the CFG microarray (see Supplementary Data 1). To further explore its specificity, we expressed O6 as a recombinant mouse Fc chimeric protein and screened the soluble form on the CFG array at three different concentrations. Within the CFG array, 63 glycans are sulfated and 4 are phosphorylated, and O6 bound in a dose-dependent manner to several glycans carrying the terminal 3-O-Sgal moiety. The highest signal intensities were observed for glycans containing the N-acetyllactosamine type II motif (Galβ1-4GlcNAc) (Fig. 1a). O6 also bound to (3 S)GalNAcβ1-4GlcNAc with a 5-fold lower signal intensity, and binding to LacNAc (Galβ1-4GlcNAc) with additional sulfates on the 4th or 6th position of the Gal either prevented or greatly reduced the signal relative to (3 S)Galβ1-4GlcNAc (Supplementary Data 1). Glycans with 3-O-Sgal and an additional 6-O-sulfate on the penultimate GlcNAc residue were bound by O6, whereas there was no recognition of sulfated terminal glucuronic acid (GlcA) and nonsulfated LacNAc glycans (Supplementary Table S1).

Notably, O6 did not bind to the glycosaminoglycan (GAG) disaccharides on the CFG array, which is a well-known class of...
glycans that carry sulfate moieties. To exclude the possibility that the VLRB was simply interacting with the presence of a negatively charged functional group, we also screened O6 on a microarray containing glycopeptides with tyrosine sulfate. O6 did not bind to the glycosulfoproteins on this array (Supplementary Data 2), further suggesting that O6 does not bind sulfated tyrosine residues, and therefore, interacts with the presence of a negatively charged functional group. The complete binding profile of O6 on the different glycans can be found in Supplementary Data 1–4.

O6 binds 3-O-SGal with low micromolar affinity. The affinity and specificity of O6 for the (3 S)Galβ1-4GlcNAc was then analyzed by isothermal titration microcalorimetry (ITC) with several sulfated and nonsulfated glycans from the CFG glycan repository. These glycans included (3 S)Galβ1-4GlcNAc, (3 S)Galβ1-3GlcNAc, (6 S)Galβ1-4GlcNAc, (3 S)Galβ1-4(Fucα1-3)GlcNAcβ-Sp, and Galβ1-4GlcNAc. Two independent experiments were performed using a high and low concentration of the disaccharides. In both experiments, curve-fitting resulted in highly similar Kd values of 10 μM (Fig. 1b). Consistent with the glycan microarray data, no binding was detected for the nonsulfated Galβ1-4GlcNAc disaccharide, isomeric β1-3-linked (3 S)Galβ1-3GlcNAc disaccharide, isomeric 6-O-sulfated (6 S)Galβ1-4GlcNAc disaccharide, or (3 S)Galβ1-4(Fucα1-3)GlcNAcβ-sparsicarboxylate (Fig. 1c). The ITC data, which is consistent with the CFG glycan microarray data, confirm that O6 binding requires a 3′-O-sulfate group, free 4′-OH and 6′-OH groups on the Gal, and a free 3′-OH group and the 2-N-acetyl group on the GlcNAc residue.

Structure of O6-glycan complex provides insights into glycan recognition. To better understand the molecular determinants of glycan recognition, we determined crystal structures of unliganded (apo) O6 and a co-crystal structure of O6 in complex with (3 S)Galβ1-4GlcNAc. This O6 VLR is relatively small compared to other characterized VLRBs, containing 178 amino acids (~19.2 kDa) and only a single LRR variable domain (LRRV1) along with the other usual LRR and CP components (LRRNT, LRR1, LRRV, CP, and LRRCT) (Fig. 2a). Apo O6 has the same overall structure as previously described VLR ectodomains with three-variable LRRs (LRR1, LRRV1, LRRV2) and a long loop in LRRCT, which extends over the concave surface, similar to other sugar-binding LRRs and is structurally most similar to the hen egg lysozyme binding VLR. Although the two do not bind the same epitope, they share 70% sequence identity and the structures align with an RMSD of 0.42 Å for 167 aligned residues (Supplementary Fig. S1A).

The structure of O6 in complex with (3 S)Galβ1-4GlcNAc at 1.9 Å resolution provides a detailed view of sulfated glycan recognition by this VLR (Fig. 2b). The sulfate group of the ligand forms hydrogen bonds with Gln86, His109, and the backbone amide of Gly137 in an identical manner to a phosphate ion in the apo structure (Fig. 2c, Supplementary Fig. S1B). Many hydrogen bonds provide specificity to the recognition of the Gal and GlcNAc moieties. Specifically, Asp85 and Ser61 hydrogen bond to the 4-OH and 6-OH of Gal, respectively. The side-chain NH of Trp35 hydrogen bonds with 3-OH of GlcNAc, while Asp16 hydrogen bonds with the nitrogen of the GlcNAc acetylamide group. Two water molecules mediate interactions between Ser37 and Asn62 and GlcNAc and Gal hydroxyl and O groups (Fig. 2c,

![Figure 1](https://example.com/figure1.png)
Supplementary Fig. S2). Trp138 of the C-terminal loop packs against the face of the GlcNAc residue, a feature that has been observed in all glycan-binding VLRs to date. The binding interactions observed in the crystal structure are consistent with the CFG glycan array and ITC data where VLR O6 makes specific interactions with 3-OH of GlcNAc and the 3-SO3, 4-OH, and 6-OH of Gal. The 2-OH of Gal, 6-OH and 1-OH of GlcNAc remain open for substitution or glycosidic linkage. Comparison of the apo and glycan complexes shows no movement of the C-terminal loop and only small changes in residue rotamers (including Trp138), indicating that the VLR is preconfigured for glycan binding (RMSD 0.28 Å for 172 aligned residues) (Supplementary Fig. S1C). The confirmation of the carbohydrate ligand is well supported by electron density and the simulated annealing OMIT map can be found in Supplementary Fig. S3.

For flow cytometry staining, we chose four CHO cell lines that have different glycomic profiles and, while it has been previously reported that GAL3ST1-4 are found within the Cricetulus griseus (Chinese hamster) genome, it has not been reported which or if these sulfotransferases are expressed in these specific cell lines. Wild-type Pro-5 cells express complex N-glycans with extended poly-LacNAc (-3Galβ1-4GlcNAcβ1-)n chains (www.functionalglycomics.org), and bind O6-mFc (Fig. 3d). Interestingly, O6-mFc did not bind to Lec8 cells, which is consistent with their lack of galactose on N- and O-glycans, as they are unable to extend the poly-LacNAc glycans due to 97% deficiency in the UDP-galactose transporter, yet are able to express normal levels of glycosaminoglycans. We also screened O6-mFc on the two engineered cell lines that express the Caenorhabditis elegans β1,4-N-acetylgalactosami- nyltransferase (Lec8GT cells) to generate the LacdiNAc (LDN) antigen (GalNAcβ1-4GlcNAc-R), and Lec8GTFT cells that also express the human α1,3-fucosyltransferase IX to generate the fucosylated LDN structure LDNF (GalNAcβ1-4GlcNAcβ1-3GlcNAc-R). O6 binds to the Lec8GT cells expressing the LDN motif; however, the addition of the fucose (LDNF) prevents recognition by the antibody. These data are consistent with and can be explained by the co-crystal structure, which suggests that addition of an N-acetyl group to the 2-OH of Gal should not notably interfere with O6 binding. Since LDN-expressing CHO cells were bound by O6-mFc, the results indicate that the GAL3ST1-4 enzymes in CHO cells are able to generate 3-O-sulfated GalNAc residues using glycans generated by the C. elegans β1,4GlcNACT and, as shown above, can be detected with O6, albeit far more weakly than wild-type Pro-5 expressing 3-O-SGal.

**Glycoprotein microarray screening of O6 reveals a wide range of glycoproteins expressing 3-O-SGal.** We further explored the ability of O6 to recognize 3-O-SGal epitopes on glycoproteins, as the detection of sulfated glycans on glycoproteins by mass spectrometry can be challenging, and its presence may have been
overlooked in previous studies. To this end, we generated a glycoprotein microarray to allow high-throughput screening of a range of glycoproteins displaying a variety of glycosylation patterns. For a complete description of how this microarray was generated and validated, see Supplementary Material and Supplementary Data 3–11. The glycoprotein microarray consisted of 35 different glycoproteins (Supplementary Data 3, 5), which were selected based on the following criteria: relative ease of access (defined as commercial availability or ease of accessibility for in-house purification), isolation from natural sources (hence, neo-glycoproteins are not included), and the presence of some well-defined glycan structures or determinants on the glycoproteins. Internal positive control glycoproteins for O6 binding on this microarray included bovine and human thyroglobulin, normal human respiratory mucin from both nondiseased (ND) and cystic fibrosis (CF) patients. These glycoproteins have well-described glycan structures or determinants on the glycoproteins. We then treated the array with PNGaseF to further confirm glycan-specific binding of O6 to these glycoproteins and determine if the sulfated glycans were on N- or O-glycans. As it was unclear if PNGaseF would be effective at removing the N-glycans from proteins printed on a nitrocellulose platform, we used binding of the plant lectin concanavalin A (ConA), which recognizes high mannose, hybrid and biantennary complex N-glycans, as a positive control. Treatment with PNGaseF reduced ConA binding, from ~3-fold to complete elimination of binding compared to a mock-treated subarray (Supplementary Data 4). This variation most likely reflects differences in the ability of PNGaseF to access glycans on the immobilized glycoproteins. PNGaseF treatment of the array also decreased binding of O6-mFc to human and bovine thyroglobulin, which cleaves α1-3 linked fucose to the LDN motif abrogates binding in Lec8GTFT cells.

We probed the glycoprotein array with multiple concentrations of O6-mFc and observed dose-dependent binding to several glycoproteins (Fig. 4a, Supplementary Data 3). The positive control glycoproteins bovine and human thyroglobulins, normal (ND) sputum mucin, and CF mucin were bound by O6-mFc as predicted, confirming the sensitivity and validity of this platform. Our screening identified several other glycoproteins, however, that have not been reported to carry 3-O-SGal, including haptoglobin and bovine fetuin. We were especially intrigued by the observation that O6-mFc bound to asialofetuin, but not normally sialylated fetuin, suggesting that terminal sialic acid residues may mask O6 recognition of 3-O-SGal. Although the glycosylation of fetuin has been studied frequently, it has not been reported to contain 3-O-SGal.

The observation that O6 binds to asialofetuin but not fetuin suggested that there some glycans on this protein may be both sialylated and sulfated, an observation that has not been previously reported. To investigate if this pattern was observed more broadly on other proteins, we treated the array with neuraminidase A, which cleaves α2-3, α2-6, and α2-8 terminal sialic acid. The efficiency of the neuraminidase treatment was validated with Sambucus nigra lectin (SNA), which binds sialylated glycans with
Mock-Treated Neuraminidase

These results demonstrate that the 3-O-SGal epitope is present in a number of human serum glycoproteins and bovine fetuin. This observation is noteworthy, as the N-glycosylation patterns of these proteins have been well-documented and, in some instances, have been reported as correlates of disease. It was particularly surprising that bovine asialofetuin was recognized by O6 (Fig. 5b), as it has been a standard for glycan analysis in multiple bands were seen in the human plasma sample following neouraminidase treatment, including transferrin, vitronectin, human fetuin, and bovine fetuin. These results demonstrate that the 3-O-SGal epitope is present in multiple glycoproteins not previously shown to contain this epitope.

Identification of 3-O-SGal on sialylated N-glycans of human serum glycoproteins and bovine fetuin. These results suggest that a number of human serum glycoproteins may express terminal glycan modification by both sialic acid and 3-O-SGal, and that sialic acid may mask binding by O6-mFc. To further investigate this phenomenon, we chose six of the glycoproteins printed on the array, as well as albumin/IgG depleted human plasma, to screen via western blot using O6-mFc (Fig. 5a).

Fig. 4 Development of a new glycoprotein microarray reveals many human serum glycoproteins carrying galactose-3-sulfate modification. a Binding of O6-mFc on the glycoprotein array at three different concentrations, with the actual slide image shown and the plot of binding, RFU relative fluorescence units, error bars = ±1 SD. b PNGaseF treatment of the glycoprotein array eliminates many positive binders, suggesting that galactose-3-sulfate modifications are carried on N-glycans. c Treatment of the glycoprotein array with neuraminidase reveals many more positive binders and are summarized in the table d) below the chart.

\[
\begin{array}{|c|c|}
\hline
\text{Glycoprotein} & \text{Binding Strength*} \\
\hline
\text{Human Thyroglobulin} & \text{Strong} \\
\text{Bovine Thyroglobulin} & \text{Strong} \\
\text{\(\alpha_1\)-macroglobulin} & \text{Strong} \\
\text{Fibronectin} & \text{Strong} \\
\text{Bovine Asialofetuin} & \text{Strong} \\
\text{Human Normal Sputum Mucin} & \text{Strong} \\
\text{Human Cystic Fibrosis Sputum Mucin} & \text{Weak} \\
\text{Human Haptoglobin} & \text{Weak} \\
\text{Human Hemopexin} & \text{Weak} \\
\text{Bovine Submaxillary Mucin} & \text{Weak} \\
\text{Bovine IgG} & \text{Weak} \\
\hline
\end{array}
\]

\*The “strong” binder classification refers to glycoproteins bound with an SNR ≥ 5.000 with only 1 and 10 \(\mu\text{g/ml}\) I6-mFc. The “weak” binder classification refers to glycoproteins bound with an SNR ≥ 5.000 with only 0.1, 1, and 10\(\mu\text{g/ml}\) 16-mFc. The “weak” binder classification refers to glycoproteins bound with an SNR ≥ 5.000 with only 0.1, 1, and 10\(\mu\text{g/ml}\) I6-mFc.

![Image of glycoprotein microarray](image-url)
3-O-SGal modification. To this end, fetuin was treated with PNGaseF, and released N-glycans were subsequently passed through a strong anion exchange column (SAX)\textsuperscript{59}. The glycans were permethylated according to the methodology developed by Khoo and Yu\textsuperscript{24} and analyzed at each stage of the enrichment (Pre-SAX, Flow-through, and SAX elution) by MALDI-TOF in both positive and negative ion mode (Fig. 5c).

As the N-glycans have been very well-characterized from fetuin, the glycans analyzed in positive ion mode (Fig. 5c, left panel) acted as a useful positive control for the efficacy of the SAX enrichment itself. As established previously, N-glycans on asialofetuin are predominately triantennary; however, some biantennary structures are also present as indicated. The flow-through, or unbound glycans to the SAX, were all neutral as expected, and the SAX-eluted fraction contained only a single structure, a triantennary glycan with a single sialic acid. This is consistent with the prior report of fetuin containing a sialic acid linked to a penultimate GlcNAc residue\textsuperscript{60}, which is resistant to neuraminidase treatment. Thus, this sialic acid-containing glycan was retained on the SAX column due to its negative charge, but was detected in positive ion mode due to the permethylation of the sample.

SAX enrichment revealed a predominant peak ($m/z = 2561.337$) consistent with a sulfate moiety on a glycan with 6 hexoses and 5 hexNAcS (6H5N). This structure was only detected in negative ion mode, and increased in intensity by four fold in the SAX-eluted sample (Fig. 5c, right panel). The data are consistent with this structure representing a triantennary glycan; however, the specific branch that contains the sulfate group could not be determined with this type of analysis. Thus, a small proportion of N-glycans of bovine fetuin express the sulfate modification recognized by O6, but are sialylated and masked in a manner that prevents recognition by O6 without removal of the sialic acid.

Immunohistochemical analysis of normal human tissues with O6. To further probe the overall expression of 3-O-SGal, we screened a broad selection of human tissues for O6 reactivity using paraffin-embedded tissue microarrays. Due to the processing and removal of the paraffin, which removes glycosphingolipids, the observed staining can be attributed only to glycoproteins. We observed O6 staining to sections from cerebellum, eye, Fallopian tube, stomach, rectum, esophagus, kidney, pituitary, prostate, thyroid, and endometrium (Fig. 6). Nonstained or very weakly
stained tissues included the adrenal gland, bladder, bones, cerebral cortex, ovary, cervix, placenta, small intestine, colon, heart, liver, lung, testis, pancreas, skin, spinal cord, spleen, striated muscle, thymus, and tonsils (Supplementary Fig. S4). Importantly, the observed staining had distinct patterns in different tissues. For example, O6 staining of goblet cells was observed primarily in the lower to middle section of the crypts, comparable to the expression profile previously reported for GAL3ST-3 in the gastrointestinal system\textsuperscript{61}. However, some individual goblet cells showed strong positive staining whereas others were negative. Another distinct pattern of O6 staining was observed in the kidney, where only cells lining the renal tubules showed staining, and is consistent with the previously reported expression profile of GAL3ST-3 (Supplementary Table S2). Given our observations on O6 binding to the glycoprotein array, it might be predicted that treatment of fixed tissue with neuraminidase would unmask the presence of 3-O-SGal. Indeed, we observed increased binding of O6 in several tissues treated with neuraminidase, including kidney, pituitary, and prostate (Supplementary Figs. S4, S5). These results reveal the widespread occurrence of 3-O-SGal modifications and masking by

---

**Fig. 6 IHC of healthy human tissues with O6.** Several tissues stained positive with O6, including cerebellum, eye, fallopian tube, G.I.-stomach, G.I.-rectum, G.I.-esophagus, kidney, pituitary, prostate, thyroid, and endometrium. The left panels depict the staining profile prior to neuraminidase treatment and the right panels are after (+) neuraminidase. Scale bars are noted on each image.
sialic acid in tissues that are known to express GAL3ST-2/3. Interestingly, there were several carcinoma and diseased specimens on the human tissue array, including breast, melanoma, thymus cancer as well as HBV\textsuperscript{+} cirrhotic liver tissue. After neuraminidase treatment, liver and melanoma tissues were O\textsuperscript{6} positive while the breast carcinoma tissue remained negative and thymoma was positive in both treated and untreated tissues (Supplementary Fig. S6). While not conclusive, these data suggest that O\textsuperscript{6} has the potential to be a cancer specific biomarker, and further exploration of these results should be pursued with more samples that include relevant clinical data.

Discussion

We have explored the expression of the 3-O-SGal epitope using a novel monoclonal antibody O\textsuperscript{6} that specifically binds to glycans carrying the 3-O-SGal modifications. Our approach is based on identifying a variable lymphocyte receptor (VLR) using lymphocytes from lampreys immunized with human type O erythrocytes. The screening technologies used here employed a YSD library to capture the desired VLR-expressing yeast on glycan microarrays\textsuperscript{39}. We expressed the VLR O\textsubscript{6} as a recombinant O\textsubscript{6}-mFc protein and demonstrated its usefulness as the first reagent to document the expression of 3-O-SGal in different human tissues and glycoproteins. While our microarray results indicate that O\textsubscript{6} can bind both 3-O-SGal and 3-O-SGALNac, the latter has not yet been described in human glycoproteins, and thus we focused on the 3-O-SGal epitope recognized by O\textsubscript{6}.

We demonstrated that O\textsubscript{6} recognizes 3-O-SGal using multiple approaches, including glycan microarray screening, ligand binding by ITC, and a crystal structure of a 3-O-SGal containing glycan with O\textsubscript{6}. To aid in this we also created a glycoprotein microarray and identified a number of glycoproteins not previously known to carry 3-O-SGal, including human serum glycoproteins and bovine glycoprotein fetuin. Thus, 3-O-SGal occurs on a number of glycoproteins not previously described by ITC. The binding interactions observed in the crystal structure were consistent with the CFG glycan array data, ITC, and glycoprotein arrays. Most notably, the crystal structure data explains the necessity of treating the samples with neuraminidase, which would remove the sialic acid from the 6th position on the galactose, where masking might occur. O\textsubscript{6} makes specific interactions with 3-OH of GlcNAc and the 3-SO\textsubscript{3}, 4-OH, and 6-OH of Gal whereas the 2-OH of Gal, and 6-OH and 1-OH of GlcNAc remain available for substitution or glycosidic linkage. There are two notable exceptions: the O\textsubscript{6} recognition of 6 S(3 S)Galβ1-4(6 S)GlcNAc (CFG Glycan ID #22) and (3 S)Galβ1-4(Fucα1-3)(6 S)GlcNAc (CFG Glycan ID #219). Neither of these glycans would be expected to bind based on the co-crystal structure. We surmise that the observed binding is a result of avidity afforded by the combination of dense presentation of glycans on the array and the bivalent presentation of VLR-binding sites in the Fc-tagged construct used for screening.

Interestingly, O\textsubscript{6} binds to sulfatide (3-O-sulfogalactosylceramide) (Fig. 3b), indicating that the 3-O-SGal alone is sufficient for O\textsubscript{6} binding, but binding to the simple 3-O-SGal alone on the CFG glycan microarray was not observed (Fig. 1). This is not surprising, as the 3-O-SGal alone directly with a spacer/linker is small in size and adjacency to the slide matrix may affect its presentation to the antibody. Regardless, sulfatide and perhaps other glycolipids represents another potential ligand for O\textsubscript{6}. Given that red blood cells were used as the immunogen for generating O\textsubscript{6}, and the presence of low but detectable levels of sulfatide on human red blood cells (~0.5 mg/kg cell mass)\textsuperscript{65}, this raises the possibility that sulfatide might have been the antigen leading to this 3-O-SGal-specific VLRB. However, it is also possible that red blood cell glycoproteins could be expressing 3-O-SGal, but this has yet to be explored.

Following the historical discovery of 3-O-SGal in sulfatide\textsuperscript{11}, and as deduced by Yamakawa et al. in 1962\textsuperscript{66}, it was subsequently shown that the 3-O-SGal modification is present in human thyroglobulin by Spiro and Bhoyroo in 1988, and in other glycoproteins by others\textsuperscript{12,17–20,48,67}. The 3-O-SGal modification has been found within the motifs sulfo-3Galβ1-4(FucR1-3) GlcNAc-R (3′-sulfo-Le\textsubscript{a}) and sulfo-3Galβ1-3(FucR1-4)GlcNAc-R (3′-sulfo-Le\textsubscript{a}), which are important in various cellular functions and in disorders including cancer and inflammation. For example, the 3-O-SGal modification is important in regulating the function of integrin subunit α\textsubscript{V}\textsuperscript{69}, and in promoting interactions of cells to E-selectin\textsuperscript{66,67}. We anticipate the O\textsubscript{6} will be invaluable in future studies to identify glycoprotein glycans carrying the 3-O-SGal epitope.

There are some physiologically important endogenous glycan-binding proteins that appear to recognize 3-O-SGal in glycoprotein glycans, but this needs further exploration. For example, the macrophage mannose receptor\textsuperscript{71} and galectin-4\textsuperscript{72} can bind glycans with 3-O-SGal modifications, and we noted previously that two other galectin family members, galectin-2 and -3, exhibit enhanced recognition of glycans with terminal 3-O-SGal residues, whereas either 4- or 6-O-sulfation blocked binding\textsuperscript{73}. However, the overall expression and biological functions of 3-O-SGal have been enigmatic, as no reagents have been available to specifically identify this modification. Interestingly, complications in understanding expression of 3-O-SGal have also arisen, as it was shown that the plant lectin MAL-1, which is typically used to identify sialylated glycans in the sequence Neu5Acα2-3 Galβ4GlcNAc-R, can also bind 3-O-SGal in the sequence 3SGalβ1-4GlcNAc-R\textsuperscript{27}. With the availability of O\textsubscript{6}, it will now be possible to specifically identify these sulfate modifications in glycoconjugates.

The masking of the 3-O-SGal epitope by sialic acid has not been noted previously. The unmasking by neuraminidase treatment suggests that sialic acid is linked near to the 3-O-SGal epitope; possibly on (a) the galactose itself, e.g., 6-O-sialic acid linkage, (b) the penultimate GlcNAc residue, or (c) a closely associated branch of the glycan (see Fig. 7). These possibilities may not be exclusive and could all be relevant to unmasking. Interestingly, such a modification could be sulfated and sialylated galactose as in (a) of Fig. 7 has been reported in N-glycans of human plasma glycoproteins, although the linkages were not well defined\textsuperscript{74}.

Such a masking role of sialic acid has been generally noted in other cases, however, as it may block recognition of underlying
that target other sulfated glycan determinants for which no antibodies are available.

Methods
Antigen preparation. Peripheral blood samples were collected in heparin coated collection tubes from healthy volunteers at the Vaccine Center at Emory University, Atlanta, GA under an approved IRB for an unrelated study and stored at −80 °C. Coded samples were provided. Total leukocytes, erythrocytes, and plasma were isolated using lymphocyte separation medium (Corning). Erythrocytes were stored in Alsever’s solution (NaCl 4.2 g/L, sodium citrate·2H2O 8 g/L, citric acid·H2O 0.5 g/L, D-glucose 20.5 g/L) at 4 °C for a maximum of 1 month.

Animals, immunization, plasma, and leukocyte isolation. Petromyzon marinus larvae (8–15 cm in length, ~2–4 years in age) were collected from tributaries to Lake Michigan (Lampey Services, Ludington, MI) and housed in sand-lined aquarium tanks at 20 °C in the animal facility. Lampey husbandry and immunizations protocols were approved by the institutional animal care and use committees (IACUC) at Emory University and were strictly adhered to during experiments. Animals (n = 3) were anesthetized with MS-222 (0.1 g/L) and given intracerebral injections (n = 3) at 2-week intervals with human type O erythrocytes. Two weeks after the final injection, the animals were sacrificed with MS-222 (1 g/L) and exsanguinated. Total lampey blood was collected into 0.67× PBS with 30 mM EDTA and layered onto 55% percoll and centrifuged for 20 min at 400 × g. Total leukocytes were collected, washed three times in 0.67× PBS and stored in RNA later at −20 °C. Lampey plasma was collected and stored at 4 °C and examined for positive VLBB titers. Serial dilutions of lampey plasma were combined with 106 washed human erythrocytes into a v-bottom microwell plate. Samples were incubated for 1 h at room temperature and assayed for agglutination by tilting the plate for 2 min. The individual lampey with the highest reciprocal hemagglutination titer (1:10,000) was used for all further analysis and antibody isolation.

Discovery and expression of VLBB clone O6. VLBB clone O6 was discovered from a yeast surface display (YSD) library generated from lampey immunized with human type O erythrocytes. Briefly, after the immunization series was completed, VLBB-specific cDNA libraries were PCR amplified (2 μg total) from total leukocyte cDNA and electroporated into the EBY100 strain of Saccharomyces cerevisiae with 1 μg of the pCT-BDFN-ESO expression vector. YSD libraries were immediately transferred to SD-CAA liquid media for growth and recovery, and the size of the YSD library was determined by plating a serial dilution of the YSD onto solid SD-CAA agar plates (~5 × 106 transformants). The YSD was then grown for several days in SD-CAA selection media at a concentration of 1 × 106 yeast/mL, and archived at ~80 °C as the unenriched library. To enrich for erythrocyte-specific VLBB clones, the YSD libraries were grown in liquid SD-CAA media overnight at 30 °C, shaking at 260 RPM and then diluted 1:5 in fresh media and grown for an additional 3 h at 30 °C, shaking at 260 RPM. The YSD library was then transferred to SD-CAA induction media and grown for 18–24 h at 30 °C, shaking at 260 RPM. To determine if the induction of the VLBB expression was successful, the YSD library was labeled with an anti-Myc-488 monoclonal antibody and monitored by flow cytometry.

The YSD library was then enriched for human type O erythrocyte VLBB clones by incubating the surface proteins on red blood cells with biotinylated anti-IgG-link SULfo-NHS-LC-Biotin, Thermofisher Pierce) and incubating with the induced library for 1 h at room temperature on rotation. The yeast bound to the labeled erythrocytes were then captured on streptavidin-labeled magnetic beads (MACS sorting) and eluted directly into SD-CAA media and grown overnight. This library was further enriched by FACs using the same approach. The success of the MACs and FACs enrichment was monitored by flow cytometry using the anti-Myc-488 and SA-PE secondary reagents. After one round of MACs and FACs enrichment, the YSD library was incubated overnight on the CFG microarray, gently shaking at 4 °C in PBS with 0.05% tween and 1% BSA. Slides were washed with PBS-T (0.05% tween) to remove the unbound yeast from the array. Yeast clones that were bound to the array were transferred to solid media, sequenced and cloned into a modified pCDH mammalian expression vector containing the IgG2a mouse Fc constant region and a 6 × HIS tag. The VLBB O6-mFc chimeric protein was produced by transiently transfecting the expression vector into 293 F cells using polyethyleneimine (PEI) (1:3 of DNA:PEI) and 2.2 mM VPA. The protein was then purified from the media according to the standard methods recommended from the manufacturer of HisPur Cobalt Resin column (ThermoFisher Scientific).

Generation of a glycoprotein microarray. The Glycoprotein Microarray was generated using a panel of 35 glycoproteins described in Supplementary Data 5. The glycoprotein microarray also included bovine serum albumin (BSA) as a nonspecific protein interaction control, biotinylated BSA as a positive control for Streptavidin binding, and Streptavidin-Cy5 (Molecular Probes) as a landing light for microarray alignment. Glycoproteins, most of which were supplied as a dry powder, were suspended in phosphate-buffered saline (PBS; 6.7 mM potassium phosphate pH 7.5,
0.15 M sodium chloride) at a stock concentration of 0.5–20 mg/mL and then 0.2 µm filtered. Human IgM and human placental laminin, which were supplied as solutions, were obtained from an orbiter to be stored. The filtered stock solution was stored at −80 °C, except IgM, which was stored at 4 °C.

Immediately prior to printing, a single aliquot of each of the glycoprotein stock solutions was thawed on ice. The glycoproteins were then diluted to a final concentration of 0.1 mg/mL in PBS, and 25 µL of each 0.1 mg/mL glycoprotein sample was transferred to a 384-well polypropylene plate (source plate). 4% serum controls, 0.1 mg/mL bovine serum albumin (control for nonspecific protein binding), 10 µg/mL biotinylated BSA (technical positive control for Streptavidin binding), 10 µg/mL biotinylated BSA (technical positive control for Streptavidin probe activity), and 1 µg/mL Streptavidin–Cy5 (lancing light and alignment control for the CFP array at fluorescence channel) were also prepared, and 25 µL of these three controls were added to the source plate. The source plate was centrifuged at 1000 × g for 3 min to remove air bubbles.

Microarray printing was performed using a SciFLexARAY S11 microarray printer (Scienion). SuperNOVA nitrocellulose-coated slides in a 16-subarray format (Grace-Bio) were preincubated at 60% humidity and room temperature prior to printing. The source plate was kept at 9 °C before and during the printing process to minimize protein aggregation and precipitation. Printing was performed with Nozzles Type 3 (PDC 90) pins at room temperature and 60% humidity. After printing, the slides were incubated for an additional 1 h on the dock at room temperature and 60% humidity. Afterwards, the printed slides were incubated in a box overnight at 4 °C.

After overnight incubation, all slides were scanned with a GenePix 4300 A scanner (Molecular Devices) with 635 nm laser and Standard Red filter to detect salts and thus the printing spots. The slides were blocked with Super G Plus Preservative solution (Grace-Bio) according to the manufacturer’s recommended protocol. The blocked slides were stored at −20 °C until use.

O6 screening on CDF glycan microarray and glycoprotein microarray. Lamprey plasma was diluted (1:10) into TSM binding buffer and incubated on a glycan microarray slide containing 610 unique glycan structures (version 5.0)92. Detection of anti-glycan specific VLRBs present within the lamprey plasma was accomplished with 4C4, an anti-VLRB mouse IgG mAb93. The chimeric VLRB-mFc O6 was bound to the CFP screen array at five-fold concentration intervals (2 µg/mL, 10 µg/mL, and 50 µg/mL) to confirm dose-dependence. Alexa-Fluor 488 labeled goat anti-mouse IgG mAb (Molecular Probes) was used for detection of both lamprey plasma and the O6-mFc recombinant fusion protein.

The Glycoprotein Microarray v3 slides were removed from −20 °C and dried in a vacuum desiccator for at least 20 min. The slides were then sealed with a ProPlate 16-well chamber (Grace-Bio). The subarrays of interest were first incubated with 200 µL of TSM wash buffer (TSMWB; 20 mM Tris pH 7.4, 150 mM NaCl, 2 mM MgCl2, 2 mM CaCl2, 0.05% v/v Tween-20) for at least 30 min. All incubation steps were performed at room temperature on an orbital shaker (−50 RPM shaking). O6-mFc was diluted to the desired concentration using TSMWB supplemented with 10% w/v bovine serum albumin (BSA; Boval) and 0.1x Super G Blocking Buffer (Grace-Bio). Seventy microliters of diluted O6-mFc solution was applied to the subarray(s). The chambers were sealed with an adhesive strip and the slide incubated for 1 h. The subarrays were washed four times for 5 min each with 200 µL of the following order of solutions: TSMWB − 2x, TSMWB − 2x, TSMWB − 2x, TSMWB − 2x. The subarrays were then briefly washed four times with TSM (20 mM Tris pH 7.4, 150 mM NaCl, 2 mM MgCl2, 2 mM CaCl2). The bound O6-mFc was detected by incubation of the subarrays for 1 h in the dark with Alexa-Fluor 633-conjugated goat anti-mouse IgG (Molecular Probes) that was diluted to 5 µg/mL in final TSMWB supplemented with 1% w/v bovine serum albumin and 0.05% w/v bovine serum albumin (BSA; Boval) and 0.1x Super G Blocking Buffer (Grace-Bio). Seventy microliters of diluted BSA solution was applied to the subarray(s). The chambers were sealed with an adhesive strip and the slide incubated for 1 h at room temperature on an orbital shaker. The subarrays were then washed four times for 5 min each with neuraminidase buffer supplemented with 0.05% w/v Tween-20 at room temperature on an orbital shaker. The subarrays were then briefly washed four times with neuraminidase buffer and then four times with Milli-Q filtered water, and immediately screened with either SNA (as a control for sialic acid removal) or O6-mFc.

Glycosidase treatment of the glycoprotein microarray. Glycosidase treatments were performed immediately following subarray rehydration. For neuraminidase treatment, the rehydrated subarrays were washed three times for 5 min with neuraminidase buffer (50 mM sodium acetate pH 5.5, 4 mM CaCl2) at room temperature on an orbital shaker. *Arthrobacter ureafaciens* neuraminidase (Roche) was diluted to 1.5 U/mL final in neuraminidase buffer, and 70 µL of the 1.5 U/mL neuraminidase was applied to the subarray. As a mock-treatment control, 70 µL of neuraminidase buffer was added to another subarray. The subarrays were then sealed with an adhesive strip, and the slide was incubated for 3 h at room temperature on an orbital shaker. The subarrays were washed four times for 5 min each with neuraminidase buffer supplemented with 0.05% w/v Tween-20 at room temperature on an orbital shaker. The subarrays were then briefly washed four times with neuraminidase buffer and then four times with Milli-Q filtered water, and immediately screened with either SNA (as a control for sialic acid removal) or O6-mFc.

For PNGaseF treatment, the rehydrated subarrays were washed three times for 5 min with 1x G7 buffer (New England Biolabs) at room temperature on an orbital shaker. PNGaseF (New England Biolabs) was diluted 1:100 into G7 buffer. PNGaseF (70 µL of 500 U/µL) was then added to the subarray(s), while 70 µL 1x G7 buffer was added to a separate subarray to serve as a mock-treatment control. The slides were sealed with an adhesive strip and incubated at 37 °C in a humidified incubator for 24 h. The subarrays were washed four times for 5 min with G7 buffer before screening. The subarrays were washed three times with TSM following by four brief washes with Milli-Q-filtered water, and immediately screened with either ConA (as a control for N-glycan removal efficiency) or O6-mFc.

VLR O6 ectodomain for ITC and crystallization experiments. The VLR O6 ectodomain was synthesized with an N-terminal gp67 secretion signal and hexahistidine tag and cloned into pCMV 3. Plasmids were purified with a Macherey-Nagel Maxi Prep kit. Proteins were expressed by transient transfection in HEK293 cells using standard methods. For purification, 5 mL Ni-NTA resin was added directly to clarified culture supernatant and incubated overnight. The resin was washed directly with 100 mL wash buffer (50 mM Tris pH 8.0, 300 mM NaCl, 15 mM imidazole) followed by 15 mL elution buffer (50 mM Tris pH 8.0, 300 mM NaCl, 300 mM imidazole). The sample was concentrated and further purified by size exclusion chromatography on a Superdex 75 16/60 column in 25 mM Tris pH 8.0, 150 mM NaCl.

Isothermal titration microcalorimetry. An Auto-ITC200 instrument (GE Healthcare) was used to perform isothermal titration calorimetry (ITC). O6 recombinant protein was purified, and lyophilized glycans were resuspended, in a buffer containing 25 mM HEPES pH 7.5 and 250 mM sodium chloride. Glycans were obtained from the CFG or purchased from Glyco tech or Sigma and contained an Sp-biotin linker (CH2CH2NH-biotin). Glycans were placed in the syringe at a concentration of 1–5 mM, and O6 was placed in the cell at a concentration of 44–67.7 µM. The O6 concentration was determined by UV absorbance at 280 nm using calculated extinction coefficients. Experiments were carried out at 25 °C and consisted of 16 injections of 2.45 µL, with injection duration of 1 s, injection interval of 180 s, and reference power of 5 µcal. Fitting of integrated titration peaks was performed with Origin 7.0 software using a single-site binding model to measure the affinity constant (Kd) value. The Kd value was then calculated as the inverse of the Kd.

Crystallization and structure determination. Crystallization screening was carried out using our in-house Rigaku CrystalMation system. No crystals and very little precipitation were observed in crystallization trials with wild-type O6 ectodomain, even at concentrations as high as 50 mg/mL. The sequence of the O6 ectodomain was submitted to the surface entropy reduction prediction (SERp) server for analysis. Three sites were suggested for mutation: K18A, K80A/E81A, and Q88A/K90A. All three variants were made using standard site-directed mutagenesis protocols and expressed and purified in the same manner as the WT protein. However, the purification yield was −10-fold lower than the WT protein. Crystals of apo O6 K18A at 10 mg/mL were grown with a well solution of 1.2 M NaH2PO4, 0.8 M K2HPO4, 0.1 M glycine pH 10.5, 0.2 M LiSO4 using the sitting drop vapor diffusion method and directly flash cooled in liquid N2. Data were collected at APS 23 ID-D (GM/CA @ APS) and processed with HKL2000. Phaser was used for molecular replacement within Phenix94 using PDB ID: 1RVX as a search model. For both structure refinement and model building in COOT95 and refinement in Phenix, TLS groups were automatically identified by Phenix. Waters were manually added and edited in COOT. The final model was assessed with quality metrics within the Phenix refine interface, which utilizes MolProbity96. For both structures, the model was refined through iterative rounds of model building in COOT95 and refinement in Phenix. TLS groups were automatically identified by Phenix. Waters were manually added and edited in COOT. The final model was assessed with quality metrics within the Phenix refine interface, which utilizes MolProbity96. For both structures, the model was refined through iterative rounds of model building in COOT95 and refinement in Phenix. TLS groups were automatically identified by Phenix. Waters were manually added and edited in COOT. The final model was assessed with quality metrics within the Phenix refine interface, which utilizes MolProbity96. For both structures, the model was refined through iterative rounds of model building in COOT95 and refinement in Phenix. TLS groups were automatically identified by Phenix. Waters were manually added and edited in COOT. The final model was assessed with quality metrics within the Phenix refine interface, which utilizes MolProbity96.
the porcine gangliosides and sulfatide was verified by staining dried TLC plates with 0.1% orcinol in 5% sulfuric acid. The plates were immediately warmed on a hot plate until the sugars could be visualized (Supplementary Fig. S12). To detect O6 binding to the lipid compounds, after separation, the dried chromatograms were incubated for 90 seconds in a 0.5% (v/v) solution of poly-isobutyl methacrylate beads dissolved in hexane. The TLC plates were dried again, and then directly blocked with 5% dried milk in PBS-tween for 30 min at room temperature. The plates were directly probed with the O6-anti-O6 antibody, and detected as described for the western blotting procedure. Complete western and lipid blots can be found in Supplementary Fig. S12.

Flow cytometry analysis of CHO cells. All of the CHO cell lines were grown in DMEM media with 10% fetal bovine serum and Pen/Strep according to standard methods. Adherent cells were collected using Trypsin with 0.05% EDTA and DMEM media with 10% fetal bovine serum and Pen/Strep. Each well was then washed with PBS, fixed, permeabilized, and stained with O6-mFc. O6-mFc was diluted to 5 µg/mL in 100% MeOH and placed at room temperature overnight until lipid blots can be found in Supplementary Fig. S12. Detection was performed with an anti-mouse IgG-HRP at 5 µg/mL and the SuperSignal ELISA Femto Student. Washes were performed using PBS with 0.5% tween (PBS-tween).

Statistics and reproducibility. Glycans displayed on the CFG glycan microarrays were printed in replicates of six. The mean relative fluorescent units (RFU) was calculated by averaging the fluorescent intensity of four median value glycan spots, which removed the highest and lowest values. From these values, the standard deviation, which is represented in the error bars in each graph, and %CV were calculated. Glycans were printed on nitrocellulose in replicates of four, and all of the values were included in the calculation of the mean RFU and standard deviation, which is represented in the error bars in each graph. The O6 VLR was run at multiple concentrations, to increase confidence in the observed binding-patterns and look for concentration-dependent binding interactions.

Flow cytometry analysis of CHO cells. All of the CHO cell lines were grown in DMEM media with 10% fetal bovine serum and Pen/Strep according to standard methods. Adherent cells were collected using Trypsin with 0.05% EDTA and DMEM media with 10% fetal bovine serum and Pen/Strep. Each well was then washed with PBS, fixed, permeabilized, and stained with O6-mFc. O6-mFc was diluted to 5 µg/mL in 100% MeOH and placed at room temperature overnight until lipid blots can be found in Supplementary Fig. S12. Detection was performed with an anti-mouse IgG-HRP at 5 µg/mL and the SuperSignal ELISA Femto Student. Washes were performed using PBS with 0.5% tween (PBS-tween).

Flow cytometry analysis of CHO cells. All of the CHO cell lines were grown in DMEM media with 10% fetal bovine serum and Pen/Strep according to standard methods. Adherent cells were collected using Trypsin with 0.05% EDTA and DMEM media with 10% fetal bovine serum and Pen/Strep. Each well was then washed with PBS, fixed, permeabilized, and stained with O6-mFc. O6-mFc was diluted to 5 µg/mL in 100% MeOH and placed at room temperature overnight until lipid blots can be found in Supplementary Fig. S12. Detection was performed with an anti-mouse IgG-HRP at 5 µg/mL and the SuperSignal ELISA Femto Student. Washes were performed using PBS with 0.5% tween (PBS-tween).

Flow cytometry analysis of CHO cells. All of the CHO cell lines were grown in DMEM media with 10% fetal bovine serum and Pen/Strep according to standard methods. Adherent cells were collected using Trypsin with 0.05% EDTA and DMEM media with 10% fetal bovine serum and Pen/Strep. Each well was then washed with PBS, fixed, permeabilized, and stained with O6-mFc. O6-mFc was diluted to 5 µg/mL in 100% MeOH and placed at room temperature overnight until lipid blots can be found in Supplementary Fig. S12. Detection was performed with an anti-mouse IgG-HRP at 5 µg/mL and the SuperSignal ELISA Femto Student. Washes were performed using PBS with 0.5% tween (PBS-tween).

Flow cytometry analysis of CHO cells. All of the CHO cell lines were grown in DMEM media with 10% fetal bovine serum and Pen/Strep according to standard methods. Adherent cells were collected using Trypsin with 0.05% EDTA and DMEM media with 10% fetal bovine serum and Pen/Strep. Each well was then washed with PBS, fixed, permeabilized, and stained with O6-mFc. O6-mFc was diluted to 5 µg/mL in 100% MeOH and placed at room temperature overnight until lipid blots can be found in Supplementary Fig. S12. Detection was performed with an anti-mouse IgG-HRP at 5 µg/mL and the SuperSignal ELISA Femto Student. Washes were performed using PBS with 0.5% tween (PBS-tween).
8. Mitsuoka, C. et al. Identification of a novel carbohydrate capping group of the L-selectin ligand on high endothelial venules in human lymph nodes as 6-sulfotially Lewis X. J. Biol. Chem. 273, 11225–11233 (1998).

12. Honke, K. & Taniguchi, N. Sulfotransferases and sulfated oligosaccharides. J. Biol. Chem. 288, 23597–23606 (2013).

13. Suzuki, A. et al. Paranodal axoglial junction is required for the maintenance of CNS myelin and axon structure. Glia 32, 166–174 (2000).

21. Chandrasekaran, E. V. et al. 6-Sulfotially Lewis X. J. Biol. Chem. 286, 22961–22968 (2011).

22. Seko, A., Hara-Kuge, S. & Yamashita, K. Molecular cloning and characterization of a novel human galactose 3-O-sulfotransferase that is specific in translocation of UDP-galactose across Golgi vesicle membranes. J. Biol. Chem. 279, 26377–26386 (2004).

27. McKitrick, T. R. et al. Antibodies from lampreys as smart anti-glycan reagents for cholangiocarcinoma. Oncotarget 8, 11094–11104 (2017).

30. Galsky, R. G. & Et et al. Epitope diversity of N-glycans from bovine peripheral myelin glycoprotein P0 revealed by mass spectrometry and nano probe magic angle spinning 1H NMR spectroscopy. J. Biol. Chem. 276, 30834–30844 (2001).

37. Hong, X. et al. Sugar-binding proteins from Leucine-rich-repetitive variable lymphocyte receptors as modules to target plant-expressed proteins. Plant Methods 13, 29 (2017).

45. Velikovsky, C. A. et al. Structure of a lamprey variable lymphocyte receptor in complex with a protein antigen. Nat. Struct. Mol. Biol. 16, 725–730 (2009).

46. Chao, G. et al. Isolating and engineering human antibodies using yeast surface display. Nat. Protoc. 7, 755–768 (2006).

57. Kim, J. H. et al. Direct analysis of aberrant glycosylation on haptoglobin in patients suffering from cystic fibrosis. J. Biol. Chem. 289, 18794–18813 (2014).

65. Green, E. D., Adelt, G., Baeniger, J. U., Wilson, S. & Van Halbeek, H. The asparagine-linked oligosaccharides on bovine fetuin. Structure of N-glycanase-released oligosaccharides by 300 megahertz 1H NMR spectroscopy. J. Biol. Chem. 263, 18253–18268 (1988).

66. Debruyne, E. N. et al. Diagnostic value of the hemepoxin N-glycan profile in hepatocellular carcinoma patients. Clin. Chem. 56, 823–831 (2010).

70. Lewis, N. E. et al. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. Nat. Biotechnol. 31, 795–769 (2013).

71. Deutscher, S. L. & Hirschberg, C. B. Mechanism of galactosylation in the Golgi apparatus. A Chinese hamster ovary cell mutant deficient in translocation of UDP-galactose across Golgi vesicle membranes. J. Biol. Chem. 269, 18794–18813 (1994).

72. Prasanthanich, N. S. et al. Immunization with recombinantly expressed glycan antigens from Schistosoma mansoni induces glycan-specific antibodies against the parasite. Glycobiology 24, 619–637 (2014).

73. Lamblin, G. et al. Structure of two sulphated oligosaccharides from respiratory mucins of a patient suffering from cystic fibrosis. A fast atom bombardment m.s. and 1H-n.m.r. spectroscopic study. Biochem. J. 275, 199–206 (1991).

74. Green, E. D., Adelt, G., Baeniger, J. U., Wilson, S. & Van Halbeek, H. The asparagine-linked oligosaccharides on bovine fetuin. Structure of N-glycanase-released oligosaccharides by 300 megahertz 1H NMR spectroscopy. J. Biol. Chem. 263, 18253–18268 (1988).

75. Debruyne, E. N. et al. Diagnostic value of the hemepoxin N-glycan profile in hepatocellular carcinoma patients. Clin. Chem. 56, 823–831 (2010).

76. Harris, C. T. et al. Fibrinogen in cell adhesion and migration via N-glycosylation. Oncotarget 8, 70653–70668 (2017).

77. Kim, J. H. et al. Direct analysis of aberrant glycosylation on haptoglobin in patients with gastric cancer. Oncotarget 8, 11094–11104 (2017).

78. Betesh, L. et al. Identification of fucosylated Fetuin-A as a potential biomarker for cholangiocarcinoma. Proteomics Clin. Appl. doi:10.1002/prca.201600141 (2017).
62. Hansson, C. G., Karlsson, K. A. & Samuelsson, B. E. The identification of sulphatides in human erythrocyte membrane and their relation to sodium-potassium dependent adenosine triphosphatase. J. Biol. Chem. 283, 813–819 (1978).

63. Yamakawa, T., Kiso, N., Handa, S., Makita, A. & Yokoyama, S. On the structure of brain cerebroside sulfonic ester and ceramide dihexoside of erythrocytes. J. Biol. Chem. 52, 226–227 (1962).

64. Capon, C. et al. Structures of O-glycosidically linked oligosaccharides isolated from human meconium glycoprotein. Eur. J. Biochem. 182, 139–152 (1989).

65. Chandrasekaran, E. V., Jain, R. K. & Matta, K. L. Ovarian cancer alpha 1,3-L-lectin to 3-sulfated Le(a)- and Le(x) oligosaccharides and the predominance of sulphate in this interaction demonstrated by binding studies with a series of lipid-linked oligosaccharides. Biochem. Biophys. Res. Commun. 188, 244–251 (1992).

66. Yoon, W. H., Park, H. D., Lim, K. & Hwang, B. D. Effect of O-glycosylated mucin on invasion and metastasis of HM7 human colon cancer cells. Biochem. Biophys. Res. Commun. 222, 694–699 (1996).

67. Yuen, W. H., Park, H. D., Lim, K. & Hwang, B. D. Effect of O-glycosylated mucin on invasion and metastasis of HM7 human colon cancer cells. Biochem. Biophys. Res. Commun. 222, 694–699 (1996).

68. Capon, C. et al. Sulfated lewis X determinants as a major structural motif in carbohydrate structures. Nat. Struct. Mol. Biol. 22, 833–834 (2015).

69. Emes, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. 66, 486–501 (2010).

70. Chen, V. B. et al. MOLPRObity: all-atom structure validation for carbohydrate crystallography. Acta Crystallogr. D. Biol. Crystallogr. 66, 12–21 (2010).

Acknowledgements

We acknowledge funding from the National Institute of Health: P41GM103694 (R.D.C.), R21GM137763 (R.D.C.), U10CA199882 (R.D.C.), R01AI072435 (M.D.C.), R01AI042266 (I.A.W.), and this project was supported by award number T32AI074492 from the National Institute of Allergy and Infectious Diseases (T.R.M.), and the Independent Research Fund Denmark (7005–00083B) (C.K.G.), and the Skaggs Institute for Chemical Biology (I.A.W.), GM/CA@APS has been funded by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006, P30GM13896). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM130399). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Author contributions

T.R.M. isolated the VLRB protein, performed microarray, western blotting, ELISA, flow cytometry, and MALDI analysis. S.M.B. expressed the VLRB ectodomain and all crystal structure experiments. A.L.N. developed the glycoprotein array, and all of the on slide enzyme treatments and lectin staining. B.C.C. performed the ITC experiments. C.K.G. performed the IHC. A.M.M. ran microarray experiments and analyses. J.H.M. contributed to the experimental design, writing, and execution of the microarray experiments. B.R.H. contributed to the experimental design. I.A.W., M.D.C., and R.D.C. are the authors and their intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-021-02199-7.

Correspondence and requests for materials should be addressed to R.D.C.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.