

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Probing Host-Microbe Interactions Through Glycomic and Glycoproteomic Methods

Permalink

<https://escholarship.org/uc/item/1d44z2c5>

Author

SHENG, YING

Publication Date

2021

Peer reviewed|Thesis/dissertation

**Probing Host-Microbe Interactions
Through Glycomic and Glycoproteomic Methods**

By

YING SHENG
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry, Molecular, Cellular and Developmental Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Prof. Carlito B. Lebrilla, Chair

Prof. David A. Mills

Prof. Andrew J. Fisher

Committee in Charge

2022

ACKNOWLEDGE

I want to take this opportunity to offer my heartfelt appreciation to my family. They always support me. They are my source of strength. I want to share my deepest gratitude to Carlito Lebrilla for being such a great mentor. He is not just a principal investigator who gives me indications regarding my projects in lab but also guides me outside the lab all these years. His wisdom and decency have been invaluable as I forge my way through the greater scientific community. When I look back on my first meeting with Carlito, I think I cannot be luckier to join this lab and know this man.

I also would like to thank all the members of the Lebrilla Lab, especially Yixuan Xie and Siyu Chen, for being such wonderful lab mates and all the lively discussions; Maurice Wong for his intellectual contribution; Diane Tu for her company with me; Armin for his help; and Winnie for bearing with me.

I sincerely appreciate my dissertation committee members, Dr. David A. Mills and Dr. Andrew J. Fisher, for revising my dissertation. I genuinely appreciate it. Thank you

Thank you to everyone who has helped me along the way. Last but not least, my profound gratitude goes to a girl who came to UC Davis for her graduate study 4 years ago. I want to thank her for not giving up and always pursuing her passion for science and enthusiasm in life.

ABSTRACT

Microbe-host interactions are mediated by protein-carbohydrate binding processes. The microbial adherence to cellular targets is a crucial step in pathogenesis. N-Glycans are oligosaccharides attached to the polypeptide of proteins and can be found on the surface of mammalian cells. Within the respiratory and intestinal tracts, highly glycosylated epithelial cells represent the primary boundary separating embedded host tissues from pathogens. Currently, there are limited methods to comprehensively explore the roles of glycans involved in these infections, therefore both virus and bacterium were employed in this study to investigate this topic. SARS-CoV-2, the causative agent for the COVID-19 pandemic, reaches into the respiratory tract, and *Salmonella*. Typhi can infect the intestinal tract and cause typhoid fever.

The cell surface glycome was manipulated via metabolic engineering. N-Glycans on the cell underwent cell membrane extraction, enzymatic release, enrichment and was eventually analyzed with an Agilent 6520 Accurate Mass Q-TOF LC/MS equipped with a PGC micro-fluidic chip. Peptides and glycopeptides from host cells were obtained using cell membrane extraction and trypsin digestion. Peptides were desalted by solid-phase extraction with C18 cartridges, whereas glycopeptides were enriched with HILIC. They were then both analyzed using an Orbitrap Fusion Lumos LC-MS/MS system. Viral infection assay and confocal microscopy were employed to explore the effect of glycome on binding with SARS-Cov-2. Bacterial adherence and invasion assays were used to assess bacterial infection capacity with different types of N-glycans dominating the host cell surfaces.

We established a cell-based model that enabled us to perform reliable structure-phenotype correlative experiments and compare the effect of individual N-glycan types in

microbial infection. Using specific inhibitors, we created host cell surfaces that were primarily fucosylated, sialylated, undecorated, or contained mainly oligomannose structures. All the glycomic cell profiles here were confirmed via Q-TOF LC/MS. Confocal microscopy helped to reinforce the notion that HMOs, such as sialylated structures, can function as decoys to prevent SARS-Cov2 infection. After modifying the host cell glycosylation, binding assays showed the spike protein had a strong affinity towards sialylated N-Glycans. Combined with molecular dynamics simulations, our data further demonstrated that the spike protein, which recognized with host receptors to initiate viral entry, preferentially bound to sialic acids in α 2-3 linkage. Bacterial adherence assay illustrated that fucosylated N-glycans on HCT116 cell surface increased significantly in the number of adhered *Salmonella*. In HCT116 cells, fucosylation was tunable, providing variable amounts of exogenous fucose. We then found that adherence of *Salmonella* is associated with the abundance of host fucosylated N-glycans. Furthermore, adherence of *Salmonella* to cells could be blocked by co-incubation with fucose or pretreatment of cells with fucosidase. The results proved that fucose residues on host cells bind with *Salmonella*.

We performed qualitative and quantitative analyses of membrane proteins from host cells. The proteomic analysis showed that the metabolic engineering did not change the abundance of membrane proteins, which indicated that host protein expression did not contribute to the altered adherence. Meanwhile, glycoproteomic analysis yielded site-specific glycopeptide information. The glycopeptides were identified and quantified using a standard glycoproteomic workflow. We found that the attached N-glycans rather than glycosylation sites were manipulated in host cells. These results highlighted the importance of glycans in host-microbe interactions.

This study also developed a lectin proximity oxidative labeling (Lectin PROXL) method to identify lectin-binding glycoproteins. The lectin was modified with a probe to produce hydroxide radicals. Thus, the lectin-recognized glycoproteins are oxidized and identified using a conventional proteomic process. All the lectin probes oxidized around 70% of glycoproteins. The specificity and sensitivity of each lectin were assessed utilizing glycomic and glycoproteomic data. Furthermore, the sialic acid- and fucose-binding lectins have higher specificity and sensitivity than other lectins. This approach provides an unprecedented perspective of lectin- glycoprotein interactions and protein networks mediated by distinct glycan types on cell membranes.

Table of Contents

Chapter 1 Introduction: Probe Host-Pathogen Interaction via MS-based Glycomics and Glycoproteomics	1
1. Basic Building Blocks of Glycans.....	1
2. Biosynthesis of Protein-Bound Glycans and Glycoengineering Approaches	4
3. Biological Roles of Glycans.....	8
3. 1. Influence of Glycosylation on Stability of Glycoproteins	10
3. 2. Influence of Glycosylation on Antibodies	10
3. 3. Influence of Glycosylation during Host-Microbe Interaction.....	11
4. Overview of Glycomic Analysis	14
4. 1. Application of Liquid Chromatography in Glycomics	15
4. 2. Application of Lectin Microarrays in Glycomics	16
4. 3. Application of Mass Spectrometry in Glycomics	17
5. MS Based Platform for Cell Glycomic and Glycoproteomic Analysis	17
5. 1. Preparation Techniques for Biological Samples	17
5. 2. Mass Spectrometry for Glycomic and Glycoproteomic Analysis	22
5. 3. Integrated Preparatory Workflow for Glycomic, Proteomic and Glycoproteomic Analysis	26
6. Conclusion	27
Reference	28
Chapter 2 Host Cell Glycocalyx Remodeling Reveals SARS-Cov-2 Spike Protein Glycomic Binding Sites	33
ABSTRACT.....	33
INTRODUCTION	35
METHODS AND MATERIALS.....	38
HMO Purification.....	38
Inhibition of HMO against SARS-CoV-2	38
Cell culture and glycocalyx remodeling treatments	39
Sample Information.....	39
Immunofluorescence.....	39
Cell membrane extraction.....	40
Enzymatic N-glycan release and purification of N-glycans	41
Glycoprotein digestion and enrichment	41
Glycomic analysis with LC-MS/MS.....	42
Glycomic data analysis	43
Glycoproteomic analysis with LC-MS/MS.....	43
Glycoproteomic data analysis	44
Molecular dynamic simulation of S protein on ACE2	44
RESULTS	45
Inhibition of virus binding by human milk oligosaccharides.....	45
Determining SARS-CoV-2 binding through variable glycocalyx expression.....	46
Molecular Dynamics Calculations of ACE2 and S Protein Interactions	50
DISCUSSION	52
CONCLUSIONS	55

ACKNOWLEDGMENT	56
AUTHOR CONTRIBUTION.....	56
FIGURES.....	57
TABLES	70
REFERENCE	71
SUPPLEMENTARY INFORMATION	76
<i>Chapter 3 Systemically Modifying Cell Glycocalyx to Understand Host-Salmonella Interaction</i>	128
ABSTRACT.....	128
INTRODUCTION	130
METHODS AND MATERIALS.....	132
Cell culture and inhibitors treatment	132
Bacterial culture	132
Bacterial Adherence and Invasion Assays	132
Cell membrane extraction.....	133
Enzymatic release and purification of N-glycans.....	133
Protein digestion	134
Glycomic analysis with LC-MS/MS.....	134
Glycoproteomic analysis with LC-MS/MS.....	135
Glycomic data analysis	135
Glycoproteomic data analysis	136
RESULTS	137
Metabolically engineering glycosylation in HCT116 cell line	137
<i>S. Typhi</i> binds to fucose residues on host cell surface	138
Proteomic quantification reveals unchanged expression of most proteins on cell surfaces.....	139
Glycoproteomic analysis yields proteins associated with glycans.....	140
DISCUSSION	142
CONCLUSION	145
FIGURES.....	146
REFERENCE	152
SUPPLEMENTARY INFORMATION	154
<i>Chapter 4 Determination of the glycoprotein specificity of lectins on cell membrane through oxidative proteomics</i>	155
ABSTRACT.....	155
INTRODUCTION	157
METHODS AND MATERIALS.....	160
Modification of lectin with Fe(III) probe.....	160
Proteins and reagents	160
Cell culture.....	160
Confocal imaging of lectin	161
Oxidation of cell surface glycoprotein	161

Cell membrane extraction.....	162
Proteins digestion and purification.....	162
Proteomic analysis using LC-MS/MS.....	162
Data Analysis	163
Cell surface N-glycomic analysis	163
RESULTS	165
Production of Lectin PROXL Probes	165
Oxidative Labeling of Glycoproteins on Cell Surfaces of PNT2 Cell Lines	165
Determination of the Relationship Between Sites of Glycosylation and Sites of Oxidation	167
Glycoprotein specificity of lectins	169
Glycoprotein-protein interactions on cell membrane are probed by lectins	173
Application of the method towards LNCaP cell line	175
DISCUSSION	177
ACKNOWLEDGMENTS	180
AUTHOR CONTRIBUTIONS.....	180
FIGURES.....	181
TABLES	186
REFERENCE	187
SUPPLEMENTARY INFORMATION	190

Chapter 1 Introduction: Probe Host-Pathogen Interaction via MS-based Glycomics and Glycoproteomics

Glycans are oligosaccharides often covalently attached to peptides or lipids. These glycans are optimally positioned as the primary molecular contacts engaged during cellular encounters with viruses, bacteria, antibodies, toxins, and other host cells. Hence glycans mediate a wide range of intercellular and host-microbe interactions. This study focuses on the role human host glycans played in bacterial and viral pathogenesis. This chapter describes the biosynthesis, functions of glycocalyx involved in host-pathogen interactions. The existing approaches for investigating their structures and roles in biological processes are then explored with the emphasis on the membrane-associated glycans in mammalian cells.

1. Basic Building Blocks of Glycans

Glycans combine monosaccharide units into branched structures whose diversity results from linear or branched α or β linkages between monosaccharides(1). The monosaccharides need to be converted into activated forms before being applied as the building blocks of glycans(2). In eukaryotes, most glycosylation occurs inside the endoplasmic reticulum (ER) and Golgi apparatus, whereas nucleotide sugars are synthesized in the cytoplasm(2).

In the biosynthesis of glycan, glycosyltransferases carry the activated forms of monosaccharides as donors and transfer saccharide moieties to acceptors such as proteins, lipids, and growing glycan chains (**Figure 1.1.**)(2). Besides

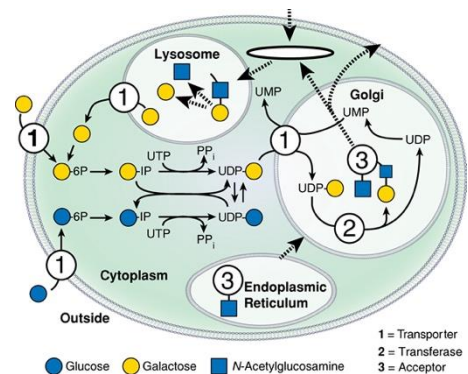


Figure 1.1. Biosynthesis, Use, and Turnover of a Common Monosaccharide.

external sugars sources, intracellular sources can also provide monosaccharides, including salvage from degraded glycans, activation, and interconversion of monosaccharides. Most degradation takes place in lysosomes, after which the monosaccharides are released and exit the lysosome. The majority of monosaccharides that reach the cytoplasm are activated and reused. The different monosaccharides commonly found in animal cells and their interconversion are shown in **Table 1.1.** and **Figure 1.2.** (2). Understanding the synthesis of building blocks, especially the control points, is helpful with glycoengineering.

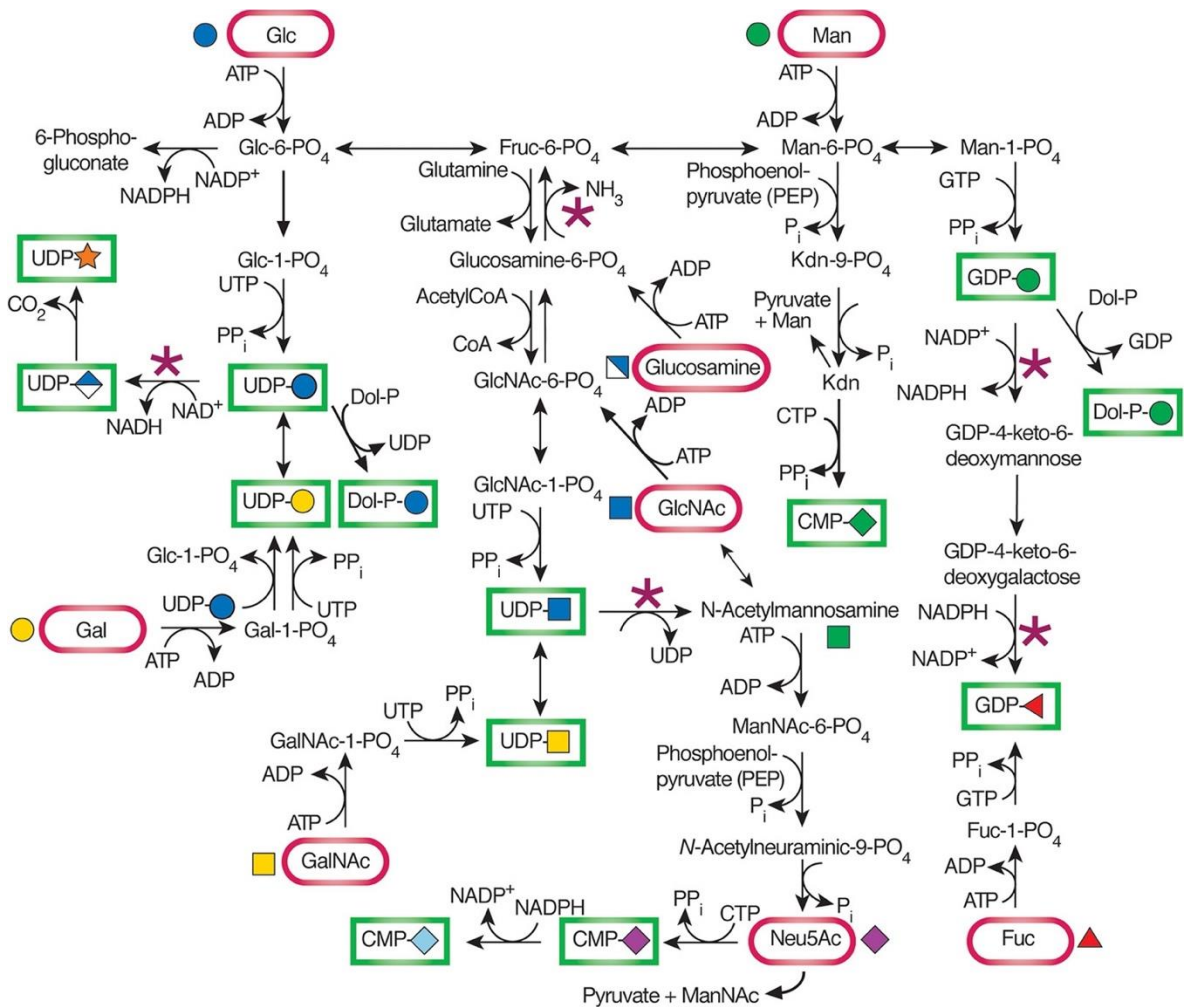












Figure 1.2. Biosynthesis and Interconversion of Monosaccharides. Rectangles) donors; (ovals) monosaccharides; (asterisks) control points; (KDN) 2-keto-3-deoxy-D-glycero-D-galactononic acid; (Dol) dolichol.

Table 1.1 Common monosaccharides: names, abbreviations, symbols, and activated forms.

Name	Abbrev	Symbol	Activated Forms
Glucose	Glc		UDP-Glc
Galactose	Gal		UDP-Gal
Mannose	Man		GDP-Man
Fucose	Fuc		GDP-Fuc
N-Acetylglucosamine	GlcNAc		UDP-GlcNAc
N-Acetylgalactosamine	GalNAc		UDP-GalNAc
N-acetylneuraminic acid	Neu5Ac		CMP-Neu5Ac
N-glycolylneuraminic acid	Neu5Gc		CMP-Neu5Gc
Xylose	Xyl		UDP-Xyl
Glucuronic Acid	GlcA		UDP-GlcA

Fucoses and NeuAcs are two important kinds of monosaccharides in human cell glycosylation. Both of their activated forms can be generated from *de novo* and *salvage* pathways(3). Mutations of the control points lead to the failure of biosynthesis. For example, NeuGc is commonly detected in mammals, and many pathogens specifically target Neu5Gc to bind and infect vertebrates. However, due to a mutation in the *CMAH* gene, which encodes the hydroxylase enzyme that converts CMP-Neu5Ac to CMP-Neu5Gc, human cells are unable to utilize Neu5Gc in glycosylation. The details of fucosylation and sialylation in human cells will be discussed more in the following few chapters, so as their significances during host-microbe interactions.

2. Biosynthesis of Protein-Bound Glycans and Glycoengineering Approaches

Based on the type of the sugar–peptide bond and the oligosaccharide connected, glycans can be classified into specific groups. The classes of eukaryotic glycans, including N-glycans, O-glycans, glycolipids, glycosylphosphatidylinositol (GPI-glycan), and C-linked glycans. The most common ones are N- and O-linked glycans. This chapter focuses on the biosynthesis of N-glycans and the common strategies to modify them, especially in mammalian cells.

N-Linked glycans bind to the carboxamido nitrogen on asparagine (Asn or N) residues in the ER, while the biosynthesis of O-glycans initiates from monosaccharides bound to the hydroxyl group of serine or threonine in the ER, Golgi, cytosol, and nucleus.

The process of N-glycosylation has been well-studied. The initiation is synthesizing Dolichol-P-P-GlcNAc₂Man₉Glc₃. To begin with, Dolichol phosphate (Dol-P) receives GlcNAc-1-P on the cytoplasmic face of the ER membrane. Dol-P-P-GlcNAc is extended to Dol-P-P-GlcNAc₂Man₅ before being “flipped” into the luminal side. More mannose and glucose residues

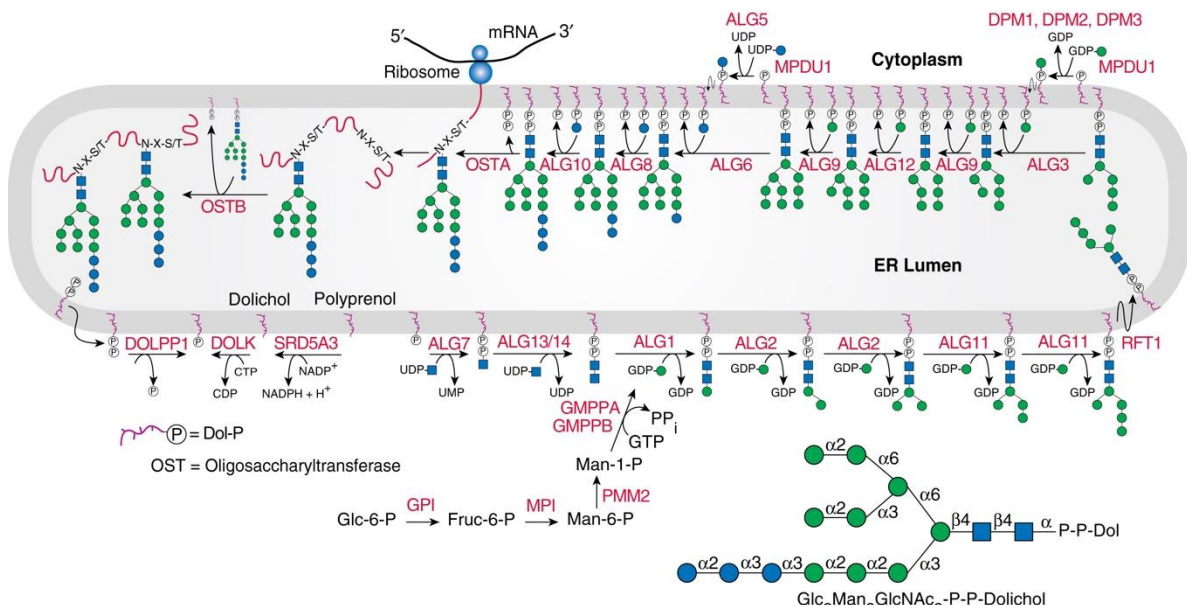


Figure 1.3. The Synthesis of Glc₃Man₉GlcNAc₂-P-P-Dol.

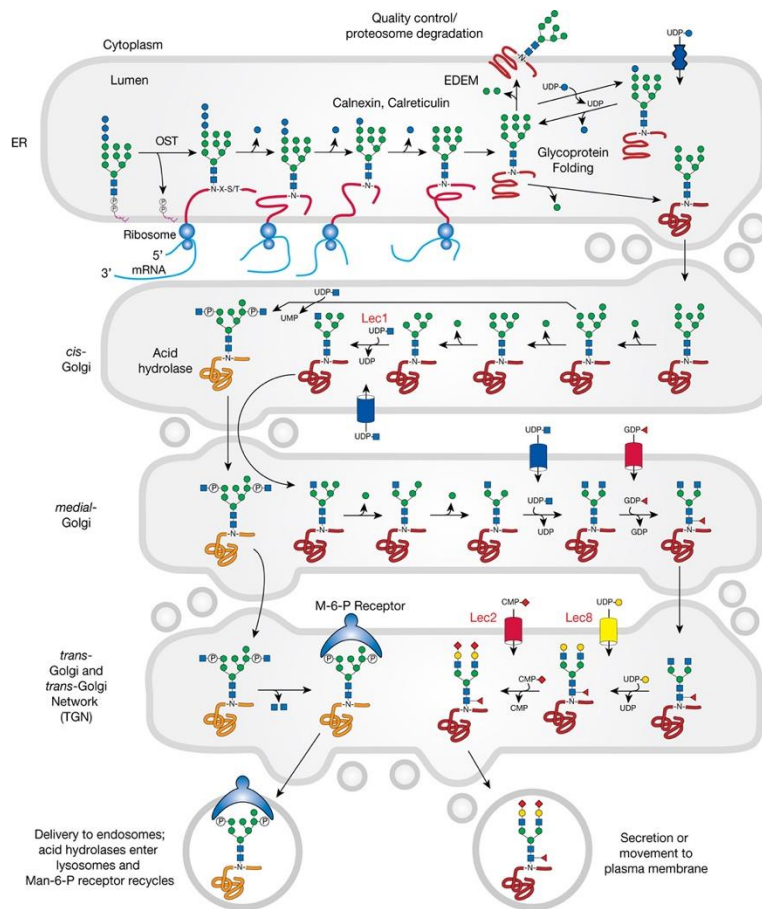


Figure 1.4. Processing and Maturation of an N-Glycan

are added to generate $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$, which is the mature N-glycan precursor (Figure 1.3.).

The second phase begins with glucose and mannose trimming, which is associated with glycoprotein folding. Subsequently, fucoses and/or sialic acids are added onto the N-glycan, as shown in Figure 1.4.. This process happens in the lumen of the ER and continues in the Golgi in a species-, cell type-, protein-, and even site-specific manner. All eukaryotic N-glycans share a common core sequence, $\text{Man}\alpha 1\text{-}3(\text{Man}\alpha 1\text{-}6)\text{Man}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-Asn-X-Ser/Thr}$.

Glycosylation has a strong influence on the biological activity of proteins. Each glycosylation has a distinct consequence. For example, $\alpha 1,6$ -fucosylation has a significant effect on ADCC (antibody-dependent-cell-mediated cytotoxicity), and $\alpha 2,6$ -sialylation is required for antibody anti-inflammatory activity(4). The biological roles of glycosylation will be discussed in

The mature glycan attached to Dolichol-P-P is usually transferred to consensus sequons (Asn-X-Ser/Thr) during protein synthesis as proteins are being translocated into the ER. Some transfers may also occur after translocation is complete.

The second phase begins

greater detail below. Due to the inability of cell lines to spontaneously manufacture their glycoforms, it is necessary to manipulate to produce proteins with optimized glycosylation based on the targeted mode of action. Thus, glycoengineering is meaningful. One of the major approaches to modify glycosylation is to apply metabolites such as metabolic glycosylation inhibitors. One of the best-known examples is the inhibitors of mannosidases. In the eukaryotic N-glycosylation pathway, mannosidases are important enzymes that are divided into two categories (I and II), each with its own catalytic mechanism, sequence, and structure. For example, kifunensine is an alkaloid with a high inhibitory effect on class I-mannosidases, specifically MAN1A1, MAN1A2, MAN1B1, and MAN1C, but a mild inhibitory effect on class II-mannosidases(5). Deoxymannojirimycin, a mannose analogue, inhibits alpha-mannosidase I in Golgi but not alpha-mannosidase in ER and leaves Man₈GlcNAc₂(6). Golgi α-mannosidase II (GMII) removes two outer Man residues, and this reaction can be blocked by inhibitor swainsonine which binds to GMII(7).

In mammals, the cores of all protein-associated glycans are conserved. However, there are substantial differences in terminal residues. Sialylation and fucosylation are often found in the termini of glycans, and their presence can also be metabolically altered. The sugar analogs can simply be added to the culture medium and enter the cells. The modified monosaccharides are processed by the biosynthetic enzymes in a manner similar to the native precursor yielding modified surface glycans(3, 4).

Metabolic oligosaccharide engineering (MOE) is a powerful tool for modifying the cell glycome. MOE reagents can function by being fed to living cells, metabolically converted into nucleotide-sugars by the cellular biosynthetic system, then incorporated into the glycome via the

activity of glycosyltransferases (GTs)(8-11). The common modifications are azides and alkynes. Cell-surface glycans bear these unnatural functional groups that can be further probed by bioorthogonal analysis. For example, azide functional groups can be introduced on sialic acids of the cell surface via the addition of per-O-acetylated azido N-acetylmannosamine (ManNAz) through the cells' biosynthetic machinery(12). The compound 6-alkynyl fucose partially inhibits fucosylation but may also be incorporated into the glycan structure. Fluorinated reagents are also widely used. The sialic acid inhibitor 2,4,7,8,9-Penta-O-acetyl-N-acetyl-3-fluoro- β -D-neuraminic acid methyl ester is activated to CMP-P-3FS and binds to sialyltransferases to decrease sialylated glycans(13, 14). Another compound, 2-deoxy-2-fluorofucose, targets fucosyltransferase VIII and inhibits the addition of fucose residues to the core of N-glycans(15). Metabolic glycoengineering is a simple technique but has advantages. The inhibitors are often highly effective. The small molecules are also easily taken up by the cells and function effectively

Genetic glycoengineering is also applied to alter glycosylation. These methods include knockdown, overexpression, and precision genome editing(2). In glycosylation, multiple glycosyltransferases are capable of producing the same linkage. For this reason, knockdown has not achieved widespread application in glycoengineering mammalian cell lines due to the limited efficacy caused by substantial and alternative glycosyltransferase activity. For example, the human fucosyltransferases III-VII all attach fucose to N-acetylactosamine moieties in α 1-3 linkage(16). Overexpression is achieved by adding the desired glycosyltransferase by transfection of glycogenes. The exogenous glycogenes are integrated into plasmid DNA with antibiotic resistance genes, and antibiotic selection is usually applied for of stable. But the instability of the introduced glycosylation traits and the use of antibiotics for selection have been problematic for

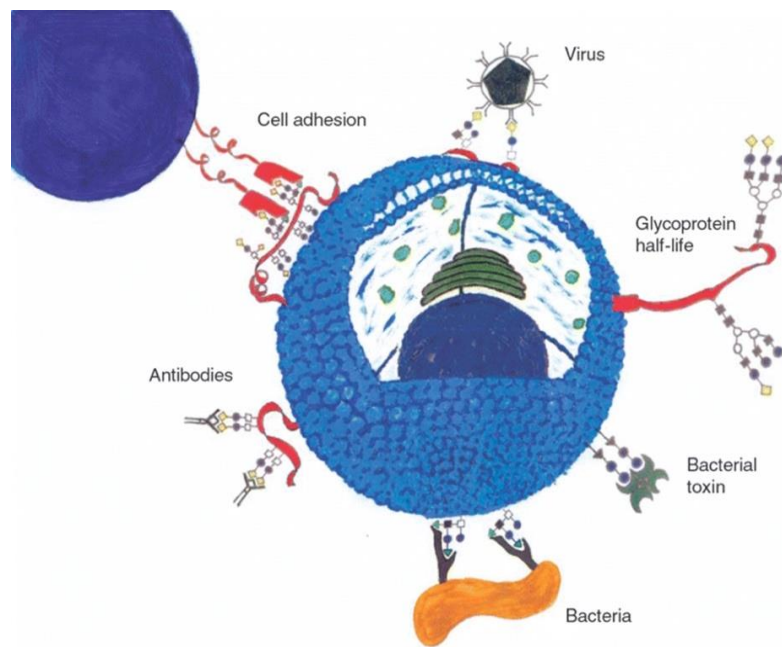


Figure 1.5. Protein–Carbohydrate Interactions at the Cell Surface

the long-term use. Precisely targeted gene editing strategies allow engineering one or more certain genes. Gene editing techniques are powerful, but knockout or knock-in strategies are currently time-consuming and difficult to use in “higher” eukaryotic cells. Due to laborious selection, these genetic

glycoengineering approaches are much less practical than metabolic inhibitors.

3. Biological Roles of Glycans

Compared to the genetic code, glycans exhibit the greatest degree of chemical complexity and evolutionary variety among distinct organisms. Glycosylation is a dense and sophisticated process that occurs in all live cells and even in the majority of viruses. Glycans are recognized as having critical metabolic, structural, and physical functions in biological systems. These carbohydrates are optimally positioned as the primary molecular contacts engaged during cellular encounters with viruses, bacteria, antibodies, toxins, and other host cells. Protein-carbohydrate interactions mediate cell-cell binding, cell–microbe (bacterial, viral, and bacterial toxin) adhesion, and cell–antibody binding (as shown in **Figure 1.5.**)(17). Sugar chains can be covalently coupled to proteins (red ribbons) or lipid-anchored in the plasma membrane. Additionally, glycosylation substitution affects the half-life of released and shed proteins in serum.

The biological roles of glycans are summarized in **Table 1.2. (18)**, while this study mainly focuses on the role of glycosylation in immunity and pathogenesis, especially its significance during host-pathogen interaction.

Table 1.2. Detailed Classification of the Biological Roles of Glycans.

Biological Roles of Glycans	
Structural and modulatory roles	Extrinsic (interspecies) recognition of glycans
Physical structure	Bacterial, fungal, and parasite adhesins
Physical protection and tissue elasticity	Viral agglutinins
Water solubility of macromolecules	Bacterial and plant toxins
Lubrication	Soluble host proteins that recognize pathogens
Physical expulsion of pathogens	Pathogen glycosidases
Diffusion barriers	Host decoys
Glycoprotein folding	Herd immunity
Protection from proteases	Pathogen-associated molecular patterns
Modulation of membrane receptor signaling	Immune modulation of host by symbiont/parasite
Membrane organization	Antigen recognition, uptake, and processing
Antiadhesive action	Bacteriophage recognition of glycan targets
Depot functions	Intrinsic (intraspecies) recognition of glycans
Nutritional storage	Intracellular glycoprotein folding and degradation
Gradient generation	Intracellular glycoprotein trafficking
Extracellular matrix organization	Triggering of endocytosis and phagocytosis
Protection from immune recognition	Intercellular signaling
Effects of glycan branching on glycoprotein function	Intercellular adhesion
Cell surface glycan: lectin-based lattices	Cell–matrix interactions
Tuning a range of function	Fertilization and reproduction
Molecular functional switching	Clearance of damaged glycoconjugates and cells
Epigenetic histone modifications	Glycans as clearance receptors
Modulation of transmembrane receptor spatial organization and function	Danger-associated molecular patterns
Masking or modification of ligands for glycan-binding proteins	Self-associated molecular patterns
	Antigenic epitopes
	Xeno-autoantigens
	Molecular mimicry of host glycans
	Convergent evolution of host-like glycans
	Appropriation of host glycans

3. 1. Influence of Glycosylation on Stability of Glycoproteins

Glycans can modify the intrinsic properties of the proteins. By hiding hydrophobic amino acids or motifs prone to aggregation, glycosylation can increase the protein's solubility and minimize the aggregation-oligomerization phenomena. Additionally, glycosylation protects these proteins from proteolysis and thermal denaturation(19, 20).

3. 2. Influence of Glycosylation on Antibodies

The majority of antibodies are glycoproteins, and their glycosylation affects their biological action, pharmacokinetics, and efficacy(21, 22). For example, the glycosylation of immunoglobulin G (fragment crystallizable) Fc domains is a key factor in antibody function diversity(23). Indeed, the N-glycan in the Fc region alters the protein's binding to Fc receptors(2), and hence its activity and half-life in circulation. Another region in the antibody, the fragment antigen-binding (Fab) region, usually carries one or more glycosylation sites and plays a role in the modulation of immune response(24). Glycans from nonhuman sources may also elicit immunogenic responses.

G0F (Hex₃HexNAc₄Fuc₁), G1F(Hex₄HexNAc₄Fuc₁), and G2F(Hex₅HexNAc₄Fuc₁) are the most abundant glycans in human antibodies(25). G0F, G1F, and G2F are abbreviations used to indicate differences in monosaccharide composition. These compounds have a core fucosylation and 0-2 terminal galactose residues. Occasionally, a terminal sialic acid is present as well(26). Although oligomannose species are frequently detected in mammalian proteins, they are present at a relatively low level in normal human antibodies(27).

Glycosylation's effect on antibody characteristics is now better understood, and glycoengineering is routinely used to improve the pharmacokinetics, affinity, and stability of

monoclonal antibodies. The effect of the glycosylation pattern on the characteristics of a monoclonal antibody can be summarized(28) as follows: nonhuman sialic acids and alpha Gals are immunogenic; sialylation enhances bioavailability; galactosylation is significant for bioactivity; both defucosylation and bisecting GlcNAc amplify ADCC-activity.

3. 3. Influence of Glycosylation during Host-Microbe Interaction

Host-microbe interactions are mediated by protein-carbohydrate binding processes(29). For decades, researchers have also investigated microbial binding to host glycans. The microbial adherence to cellular targets is a crucial step in pathogenesis. Highly glycosylated epithelial cells represent the primary boundary separating embedded host tissues from pathogens within the respiratory and intestinal tracts(30). Various pathogens express very specialized glycans on their exterior, which affects the antigenicity. Conversely, microorganisms produce highly specific exo- and endoglycosidases capable of degrading host glycans. Indeed, many of the microbial glycosidases are applied as tools to explore structural features of eukaryotic glycans(18). This study mainly focuses on the host-pathogen interaction.

3. 3. 1. Host Glycan

Several glycans are targeted by numerous viruses, bacteria, and parasites, toxins, as well as act as decoys for these constituents. The initial glycan interaction is less necessary in static situations with long contact times. However, chances for contact and infection are rare and, in reality, temporary (18). So the glycan-dependent process is significant. Given the fast development of diseases and continuous selection, the sequences of the glycans involved often exhibit excellent recognition specificity.

Microbial adhesins, including the bacterial, parasite, and fungal ones, recognize and bind with host glycans before initiating infection. One of the common examples is *Helicobacter* recognition of gastric sialylated glycans(31). Moreover, viral agglutinins can bind with host glycans. One of the most better known is influenza hemagglutinin (the “H” in “H1N1”). Their specificity toward binding the sialic acid ligand is required for infection. Additionally, many bacterial toxins are soluble and mediate their effects by binding to target glycans on host cells. For instance, toxins bind with to certain gangliosides with high specificity (*e.g.*, cholera toxin from *Vibrio cholerae* and heat-labile enterotoxin from *Escherichia coli* target GM1, shiga toxin from *Shigella dysenteriae* recognizes Gb3)(32).

The role of structural specificity is significant in glycan function. Many of the microbial binding proteins involved have been used as molecular probes to examine glycan expression. Meanwhile, some microbes have evolved the ability to hide or change glycans identified by bacteria or toxins. Many pathogens produce glycosidases to modify or destroy the host glycocalyx in the host defense system in order to favor the survival and persistence of pathogens(33) or employ the released monosaccharides as a nutritional resource(34).

Meanwhile, glycans can act as decoys to bind with pathogenic organisms or toxins before pathogenic access to host cells and be washed away. For example, the red blood cell erythrocytes use their cell surface N-glycans as receptors to bind with influenza viruses in the bloodstream and allow antigenic identification(35, 36). Other well-known decoys are mucins on the surface of epithelial cells. It is known that the transmembrane protein mucin 1 (MUC1) releases its extracellular domain to the external environment, and this domain act as a decoy toward mucosal pathogens(37)(*e.g.*, *Haemophilus influenzae*(38), respiratory syncytial virus(RSV)(39),

and *Pseudomonas aeruginosa*(40)). These decoys play a key function by specific binding with pathogens, keeping them away from their targets on the host cells(41-44).

Extrinsic glycan-binding proteins include pathogenic microbial adhesins, agglutinins, or toxins, but the pathogens also develop glycosidases to alter the colonized host glycans to evade the immune response and facilitate the infection(33). Neuraminidase (NA) has been reported significant in both viral and bacterial infections. Influenza virus NA removes terminal sialic acids of viral binding receptors in infected hosts. This enzymatical activity is essential. Infected cells do not release virions unless sialic acid residues are cleaved from the cell surface(45). Bacterial neuraminidase NanA contributes to *S. pneumoniae* adherence, desialylates monocyte surface, and promotes the crossing of the blood-brain barrier(46).

3. 3. 2. Microbial Glycan

Microbial glycans also mediate host defense. Most microbes are targets of the host's immune response, and glycan recognition is a key step in pathogenesis. A common interaction involves soluble host proteins recognizing pathogens via microbial glycans. Vertebrates produce toxic glycan-recognizing peptides capable of attacking bacteria. The antibacterial lectin RegIII γ secreted from the host small intestinal mucus layer is a typical example(47). Additionally, host galectins play functional roles in host defense by recognizing bacterial surface glycans(48). The immune cells' recognition of foreign glycans and/or pathogen-specific glycans is called pathogen-associated molecular patterns (PAMPs)(18). These microbial glycans are detected by certain receptors such as pattern recognition receptors (PRRs)(49, 50), Toll-like receptors (TLRs)(51-54), NOD-like receptors (NLRs)(55, 56), and C-type lectins(57). The common

glycans evolved in PAMPs include bacterial lipo-oligosaccharides, lipopolysaccharides, and bacterial peptidoglycans.

Meanwhile, microbes have evolved molecular mimicry of host glycans. Microbial pathogens are capable of decorating themselves with identical or similar glycans to their hosts. Glycans like these prevent recognition of antigenic epitopes, hinder triggering immune responses, and imitate host "self-associated molecular patterns" (SAMPs)(58). All the strategies outlined above can effectively evade host immune response. A well-studied example in glycan-based SAMPs is that Sia-expressing bacteria group B *Streptococcus* (GBS) binds with human Siglecs (Sialic acid-binding Ig-like lectins)-9 (hSiglec-9) to suppress neutrophil activation and facilitate bacterial survival(59).

3. 3. 3. A Novel Tool for Studying Glycan-Mediated Host-Pathogen Interactions

There are novel deep-learning resources for studying glycan-mediated host-microbe interactions(60). A dataset of species-specific glycan sequences has been recently created to facilitate the study of glycans in host-microbe interactions. This resource could be used to train machine-learning models to examine the role of glycans in host-microbe interactions. The information in glycans can predict immunogenicity, pathogenicity, and taxonomic origin. It provides a novel insight into bacterial virulence

4. Overview of Glycomic Analysis

Glycomics is a subset of glycobiology that tries to determine the structure and function of the entire set of glycans synthesized in a cell or organism, as well as all the genes that code for the enzymes involved in the glycan synthesis. A variety of approaches for glycomic analysis are currently available. Glycans can be characterized based on chromatographic retention time,

affinity (*i.e.*, lectin microarrays), and mass/charge-based glycan identification(61, 62). Researchers are also increasingly combining these techniques to obtain higher accuracy and structural insights.

4. 1. Application of Liquid Chromatography in Glycomics

A glycome often consists of a variety of glycans that are comparable in size, shape, and hydrophilicity. Liquid chromatography utilizes retention times to generate structural information(61). This utility is suited to the investigation of certain glycan populations, but it is highly dependent on the comprehensiveness and reliability of databases. **Figure 1.6.** depicts a typical LC-based glycomic analysis workflow(61). The method for preparing glycomes for LC is

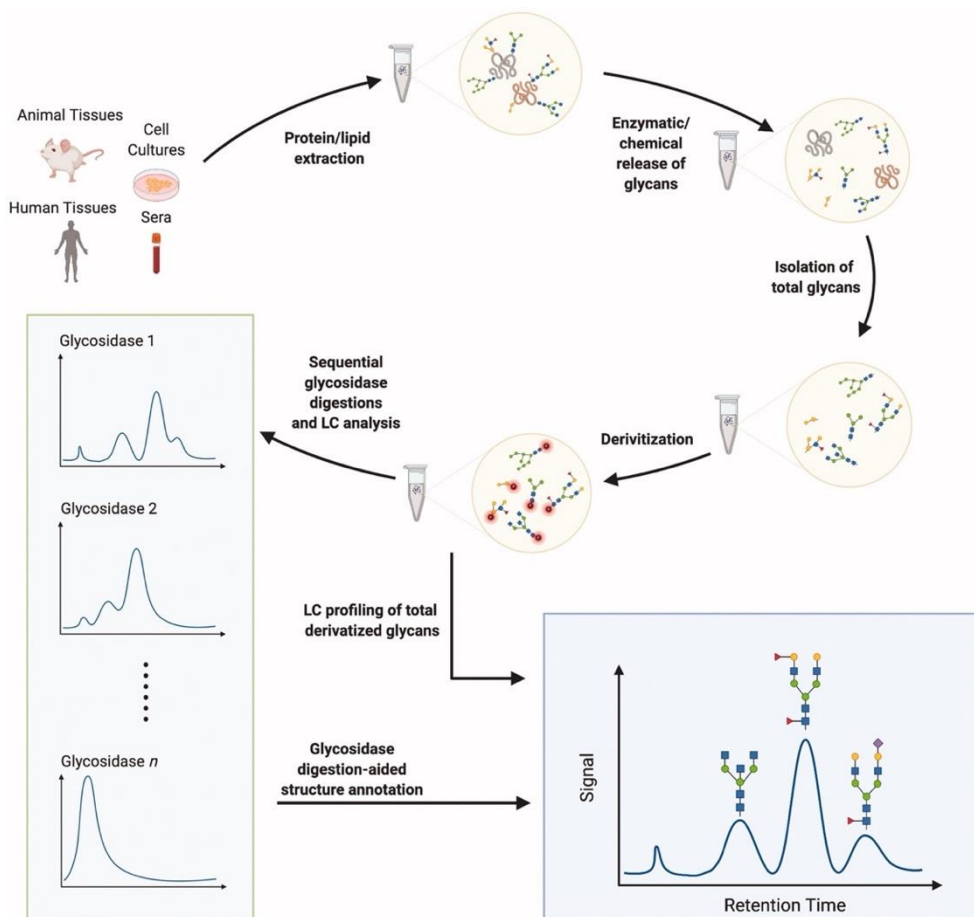


Figure 1.6. A simplified workflow of liquid chromatography-based glycomic analysis. similar to that for mass spectrometry. In order to facilitate glycan identification in LC, released

glycans are usually labeled with fluorescent tags, such as 2-aminobenzamide (2-AB) and anthranilic acid (2-AA). The quantitative analytical capacity of LC glycomics is enhanced by the 1:1 stoichiometry of the labeling reaction. In LC-based glycomics, exoglycosidase digestion is essential for structural validation. Structures are assigned based on changes in glucose units (GU) value.

4. 2. Application of Lectin Microarrays in Glycomics

Lectins are non-antibody glycan-binding proteins. A panel of lectins with various binding specificities is immobilized on well-defined support to create lectin microarrays for affinity-based identification of glycans. Compared to other approaches, lectin microarrays have several advantages. Sample preparation is relatively simple. **Figure 1.7.** depicts a single color lectin microarray for analytical purposes(61). Lectins are immobilized onto a slide using amine coupling chemistry. Samples are labelled with fluorophore tag on the lysines or amines on intact glycoproteins, followed by incubation with the lectin microarrays.

The labelling eliminates the need for glycan release or purification. After labeling, the lectins can be

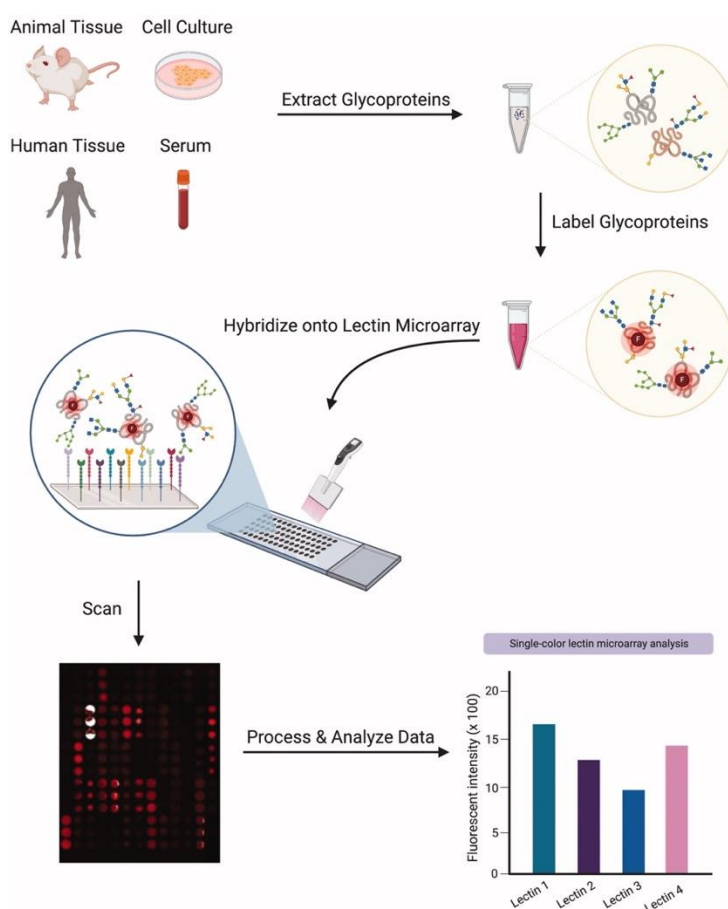


Figure 1.7. A workflow of single-color lectin microarray.

employed as probes to target glycoproteins with specific glycan signatures(63). However, the analytical capacity of lectin microarrays relies on the variety of lectins, as well as their specificities. In addition, this approach is less sensitive to glycosylation alterations(64). Pilobello *et al.* developed a dual-color lectin microarray to calibrate glycan expression (65).

4. 3. Application of Mass Spectrometry in Glycomics

The general procedure for MS-based glycomic analysis begins with sample preparation, MS analysis for structural annotation and quantitation(66, 67). MS-based analysis of glycans and site-specific glycopeptide information from glycoengineered cells will be the major approach employed in this research. The details of the workflow will be described later in this chapter.

5. MS Based Platform for Cell Glycomic and Glycoproteomic Analysis

5. 1. Preparation Techniques for Biological Samples

5. 1. 1. Cell Membrane Extraction

To fully characterize the cell surface glycoalyx, specific and effective methods for separating cell membrane from cell lysates are required. Our lab has developed a method based on density gradient centrifugation to extract cell membranes (**Figure 1.8.**)(67). To prevent biomolecule destruction, lysis is performed in a homogenization solution comprising buffers and protease inhibitors. Crude cellular fractionation can be accomplished through a series of increasing-speed centrifugation operations that separate cellular components differ in size and density. The crude membranes collected from lysis contain plasma membrane (PM), endoplasmic reticulum (ER), and plasma membrane-associated membranes (PAMs), which are microdomains of the PM interacting with the ER and mitochondria(68). The common media used for differential

centrifugation includes sucrose, mannitol, glycerol, Ficoll 400 (a sucrose polymer), Percoll (a type of colloidal silica), and iodixanol (OptiPrep). We use a discontinuous sucrose gradient to fractionate the plasma membrane. The PM can be harvested at the 43/53% sucrose interface(68). In this approach, the enriched membrane is then purified using sodium carbonate. The high alkalinity of sodium carbonate solution can dissolve the PAMs, while cell plasma membrane proteins, such as integral membrane proteins, are still insoluble in the solution(69, 70). The pellet is then washed with deionized water to eliminate any leftover sodium carbonate and prepared for further analysis.

This study mainly focused on the glycocalyx from the cell membrane but targeting specific cell membrane glycoproteins may help the future steps. There are now a variety of strategies for enriching cell membrane targets. First, affinity enrichment based on its non-covalent interactions can be utilized, such as lectin-based method and antibody-mediated immunoaffinity. Second, glycomic metabolic labeling can also be employed as handles for bioorthogonal target enrichment. Third, biphasic separation with Triton X-114 will separate the cell plasma membrane due to its amphiphilicity nature(71, 72). In addition, the negatively charged plasma membrane is caught through electrostatic interactions with the cationic colloidal silica beads and then isolated using ultracentrifugation(73).

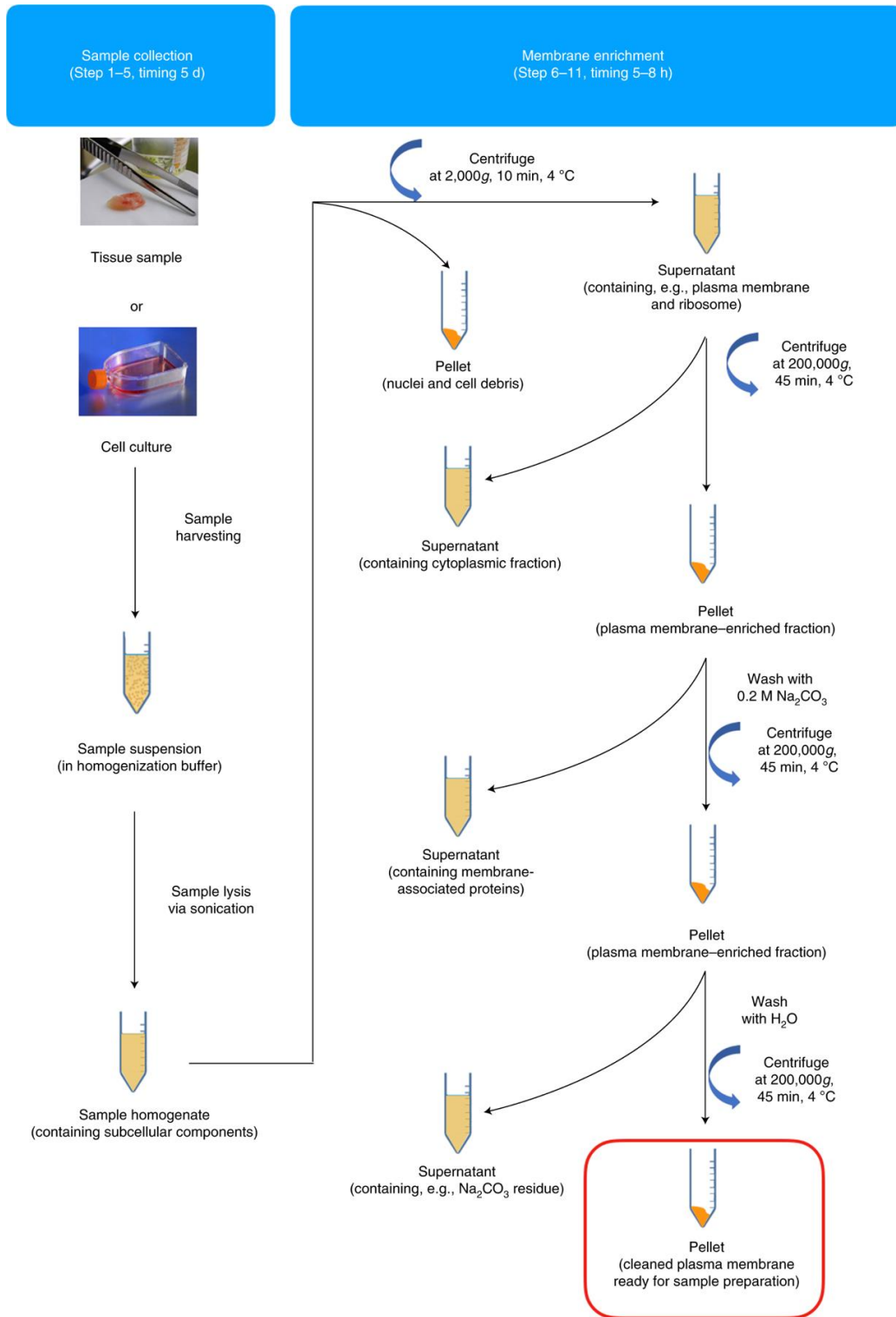


Figure 1.8. A workflow for cell sample collection and membrane extraction. At least 10 million cells are collected, then buffer exchange and sonication at 4 °C are performed. Centrifugation is employed to separate the nuclei and cell debris. After that, ultracentrifugation is used to pellet the crude plasma membrane fraction, which is then washed with aqueous sodium carbonate and water.

5. 1. 2. Release of N-Glycans

N-Glycans are released from proteins enzymatically or with chemical methods. For the analysis of N-glycans, the enzymatical release is more specific. Because they work under mild circumstances and are more convenient, endoglycosidases and glycosamidases are widely used. An amidase peptide: N-glycosidase F (PNGase F) cleaves the N-glycosidic bond between the GlcNAc and the asparagine residue, resulting in the release of N-glycans from asparagine(74, 75). Because of its broad substrate specificity and ability to hydrolyze most N-linked glycan types, PNGase F digestion is commonly used. PNGase F works for all N-glycans except those with α (1,3)-linked core fucose(76). PNGase A can cleave N-glycans with or without the core fucose residue, but its efficiency in cleaving N-glycans from glycoproteins is lower(77, 78). Additionally, there are several endoglycosidases that cleave between the two GlcNAcs in the N-glycan core regions, leaving one GlcNAc attached to the protein. Endoglycosidases, unlike PNGases F and A, exhibit selectivity for multiple types of glycans.

5. 1. 3. Protein Digestion

Glycomics data alone can characterize glycan structures, while glycoproteomics provides data regarding carrier proteins and glycosylation sites of the glycans. In order to determine the site-specific glycosylation of membrane proteins in a given cell or tissue, the glycoproteomic analysis needs to be performed on extracted cell membranes. For more efficient digestion, extracted membrane proteins are first to be dissolved with urea(79). The addition of DTT and alkylation with iodoacetamide (IAA) is required to break S-S bonds. The samples are digested thereafter with trypsin for 18 hours at 37 °C. Other enzymes with varying specificities, such as

Glu-C(80) and Lys-C(81), can be employed in tandem with trypsin to produce more extensive glycosylation site-specific mapping data.

5. 1. 4. *Solid Phase Extraction (SPE)*

5. 1. 4. 1. *Enrichment of Released Glycans*

Separation of the glycan from salts and proteins is required. Salt in the solution reduces the ionization efficiency of the target analyte, and proteins ionize more efficiently than glycans. SPE can address these problems. Due to the significant number of hydroxyl groups, glycans can be purified from other more hydrophobic compounds with hydrophilic interaction liquid chromatography (HILIC)(82). Another common SPE technique is porous graphitic carbon (PGC). In the PGC cartridge, flat sheets of hexagonally organized carbon atoms make up the packing material. The retention of glycans on PGC is influenced by a number of variables. Adsorption (solvent-solvent interactions) is one sort of interaction driven by hydrophobicity. Electronic forces also contribute to interaction(83). So PGC can be utilized to retain glycans, and PGC cartridges are now widely employed to desalt and enrich glycans prior to MS analysis. PGC employing LC-MS platforms allows for the effective separation of isomeric molecules(84).

5. 1. 4. 2. *Enrichment of Glycopeptides*

HILIC enrichment facilitates the analysis of intact glycopeptides. Due to ion suppression from the more ionizable peptides, glycopeptide analysis needs enrichment compared to peptide analysis(67). It has been reported that HILIC enrichment yields 90% glycopeptides and some residual nonglycosylated peptides (67). This enrichment process involves hydrophilic interactions between the glycans and the stationary phase. Besides HILIC, glycopeptide enrichment can be

done through other enrichment methods, including lectin affinity chromatography(85) and strong anion exchange solid-phase extraction (SAX-ERLIC)(86).

5. 2. Mass Spectrometry for Glycomic and Glycoproteomic Analysis

5. 2. 1. A Brief Introduction to Mass Spectrometry

Mass spectrometers precisely measure the molecular masses of compounds by ionizing them in the gas phase and then subjecting them to regulated electric and magnetic fields in a vacuum. Significant progress has been made in developing mass spectrometers with a large dynamic linear range and increased detection limits. Metabolomics, proteomics, glycoproteomics, and glycomics have all profited substantially from these improvements. But large molecules like proteins and glycans are chemically and thermally labile and are generally difficult to be volatilized. "Soft" ionization techniques, such as matrix-assisted laser desorption (MALDI) and electrospray ionization (ESI), allowed MS to examine biomolecules(84). ESI developed by Fenn *et al.* (87) revolutionized mass spectrometry. Dr. John Bennett Fenn received the 2002 Nobel Prize in Chemistry for this contribution. Because ESI has higher sensitivity, delivers the ion less internal energy, and is more compatible with liquid chromatography. ESI is more commonly employed to analyze glycans and glycopeptides(84). This study chose ESI instead of other ionization methods.

Great advances in mass analyzers have also been made. Each analyzer has its own operating principle and a varied mass resolution, scan rate, duty cycle, and linear range. A typical MS/MS configuration has two mass spectrometers linked by a reaction cell. The first spectrometer (MS1) selects a small mass range for fragmentation. In the reaction cell, energy is used to fragment the targeted ions. A mass spectrometer analyzes the fragments (MS2).

Common instrument combinations include triple Quadrupole (QqQ), Quadrupole-Time-of-Flight (Q-TOF), and hybrid Orbitrap instruments with numerous fragmentation modes.

5. 2. 2. Application of Glycomic Analysis with MS

A greater mass range can be produced by ionizing analytes with numerous charges. Furthermore, combining a nanoflow LC with a nano-ESI source increases sensitivity thereby enhancing the identification of glycans with low abundances. As a result, nanoflow liquid chromatography, electrospray ionization, and time-of-flight mass spectrometry (nanoLC-ESI-TOF-MS) are widely used as an analytical platform for biomarker identification and functional research. In the next subsection, more applications are discussed in greater depth. The most common way for glycans to be fragmented is through Collision-Induced Dissociation (CID), in which the analyte ions are accelerated and forced to collide with a neutral gas like nitrogen or argon. During CID, the kinetic energy of accelerated ions is converted into internal energy, leading to bond breakage and generating fragments for glycans. Data Dependent Acquisition (DDA) can pick ions based on a precursor scan. In this study, the instrument is programmed to fragment a specific number of the most abundant ions per cycle. Specific fragmentation can be set up when the precursor ion m/z is known.

5. 2. 3. Application of Glycoproteomic Analysis with MS

The identification of both the glycosylation site and the glycan structure is required for glycoprotein site-specific analysis. With contemporary mass spectrometry and data analysis methods, we can now better solve this problem. Nanoflow LC linked to hybrid Orbitrap MS is the gold standard for MS-based proteomics and is specially adapted to solve the difficulty of site-specific analysis of glycoproteins(67, 70). Orbitraps is the best-suited instrument for the

characterization of glycans and glycopeptides due to its rapid full scan rates, comprehensive coverage of fragments, and great resolving power.

For accurate identification, the MS/MS data must cover the amino acid sequence in sufficient detail and include glycan fragmentation. Peptide bonds require more energy for successful fragmentation than glycosidic bonds. Stepped collision energy can be employed to improve coverage of both glycan and peptide fragmentation. High-energy collisional dissociation (HCD) produces more fragmentations on the peptide backbones than CID. Other fragmentation mechanisms, such as electron transfer dissociation (ETD) and ultraviolet photodissociation (UVPD), can more efficiently break certain bonds and add to the MS/MS data. ETD involves transferring electrons to analyte ions using a radical or ionic reagent such as fluoranthene anions. To improve sequence coverage, high-energy collisional dissociation can be paired with electron transfer dissociation (ETHcD). Caval *et al.* have proved that increasing the mass range in ETHcD can considerably improve N-glycopeptide identification certainty(88). But ETHcD requires extended cycle time to transport the precursor ions, leading to fewer glycopeptides identified(70, 89). **Figure 1.9(90)** shows CID, HCD, and ETD fragmentation spectra of a plant glycopeptide. Minimal peptide backbone information (b6 and y2 ions) generated from the poor peptide and glycan fragmentation occurs with CID. HCD provides glycan signatures (*e.g.*, GlcNAc ion (204.09 m/z)), as well as the mass of the peptide without the glycan structure (1137.54 m/z [M + 2H]²⁺). Based on this information and MS spectra, the mass of glycan and its corresponding structure is measurable. The similarity shared by CID and HCD is the little information of peptide backbone provided. Compared to CID and HCD, ETD fragments peptide backbone, and the spectra display

most z-and c-ions. ETD reveals the confirmation of the peptide sequence and the glycosylation sites (Asn)(90, 91).

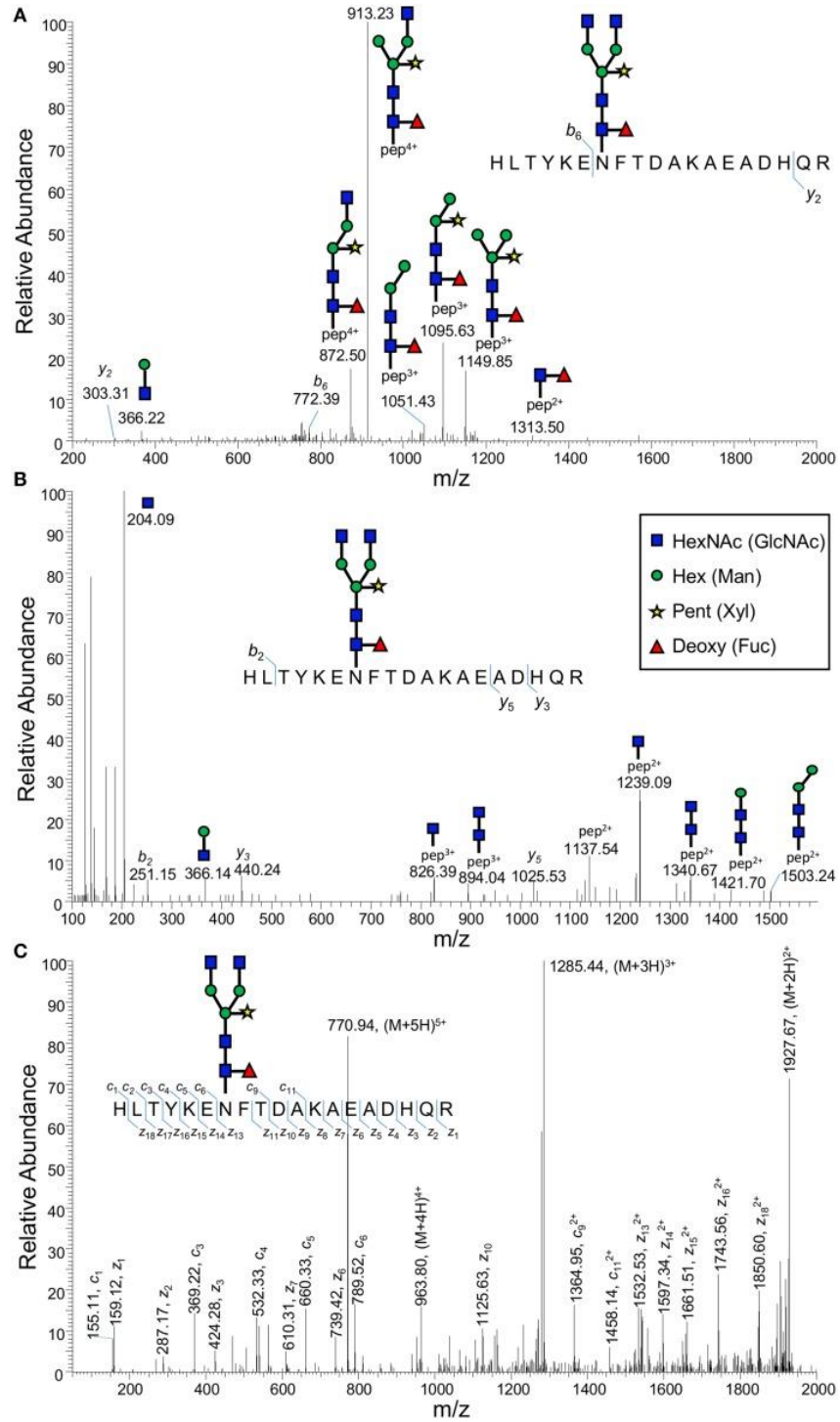


Figure 1.9. Examples of spectra from different fragmentation of glycopeptide under CID(A), HCD(B), and ETD (C), and ETHcD (right).

The massive volume of data generated by (glyco)proteomics analysis necessitates the use of automated identification methods. The amino acid sequences must be known in order to link peptide sequences to their proteins. Peptide sequences from digestion with a specific proteolytic enzyme, such as trypsin, are predictable *in silico*, and their precursor weights and fragmentation patterns can be matched with MS data using this knowledge. These estimates can be adjusted to account for natural changes like oxidation and phosphorylation, as well as synthetic alterations and known glycan compositions. There are now various software tools available that can identify (glyco)peptides. Byonic from Protein Metrics(92), MaxQuant(93), and GlycoMaster(94) are some examples.

5. 3. Integrated Preparatory Workflow for Glycomic, Proteomic and Glycoproteomic Analysis

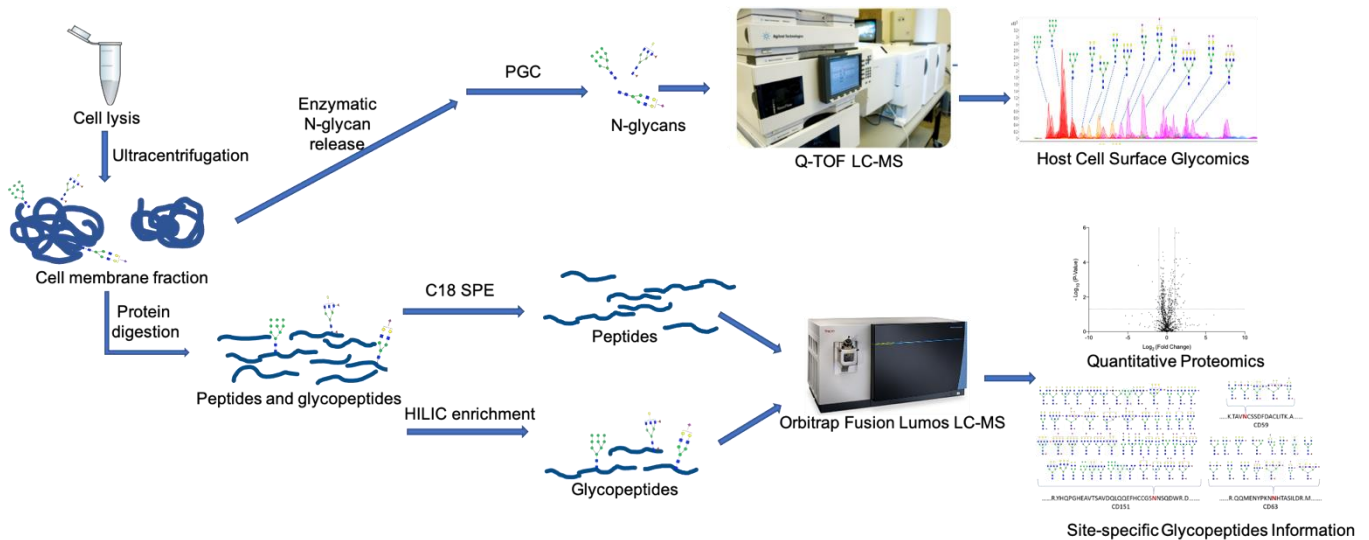


Figure 1.10. Integrated Preparatory Workflow for Glycomic, Proteomic and Glycoproteomic Analysis

The general workflow for glycomic, proteomic, and glycoproteomic analysis is shown in **Figure 1. 9.** In order to monitor the glycan profile of the cell culture, the cells were lysed and ultra-centrifuged to extract the cell membrane fraction. We obtained the cell surface glycomic

data using Q-TOF LC-MS. Simultaneously, the alterations of proteomic and site-specific glycopeptide information were monitored with an Orbitrap LC-MS platform.

6. Conclusion

Host-microbe interactions are often mediated by the carbohydrate-protein binding processes. Characterizing cell membrane glycans evolved in these interactions is critical to know how cells work. They interact directly with other cells and play an active part in distinguishing self from external invaders since they are located at the interface between the cell and its environment. To understand their interactions with other molecules, detailed structural knowledge is required.

Glycomics based on mass spectrometry has the ability to provide a single platform for monitoring cell membrane proteins at the same time. The study of cell surface glycans has been hampered by previous limitations in the analytical methods. A cell surface glycome technique was devised and applied to intestinal cells to characterize alterations throughout cellular transitions. The solutions will improve cell-based assays and allow a better understanding of the role of glycans during host-microbe interactions.

Reference

1. M. Cohen, A. Varki, Modulation of glycan recognition by clustered saccharide patches. *Int Rev Cell Mol Biol* **308**, 75-125 (2014).
2. A. Varki *et al.* (2015).
3. J. Du *et al.*, Metabolic glycoengineering: sialic acid and beyond. *Glycobiology* **19**, 1382-1401 (2009).
4. Q. Wang, M. J. Betenbaugh, Metabolic engineering of CHO cells to prepare glycoproteins. *Emerg Top Life Sci* **2**, 433-442 (2018).
5. F. Crea *et al.*, EZH2 inhibition: targeting the crossroad of tumor invasion and angiogenesis. *Cancer Metastasis Rev* **31**, 753-761 (2012).
6. J. Bischoff, L. Liscum, R. Kornfeld, The use of 1-deoxymannojirimycin to evaluate the role of various alpha-mannosidases in oligosaccharide processing in intact cells. *J Biol Chem* **261**, 4766-4774 (1986).
7. J. M. van den Elsen, D. A. Kuntz, D. R. Rose, Structure of Golgi alpha-mannosidase II: a target for inhibition of growth and metastasis of cancer cells. *EMBO J* **20**, 3008-3017 (2001).
8. E. M. Sletten, C. R. Bertozzi, Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew Chem Int Ed Engl* **48**, 6974-6998 (2009).
9. A. Cioce, S. A. Malaker, B. Schumann, Generating orthogonal glycosyltransferase and nucleotide sugar pairs as next-generation glycobiology tools. *Curr Opin Chem Biol* **60**, 66-78 (2021).
10. L. K. Mahal, K. J. Yarema, C. R. Bertozzi, Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis. *Science* **276**, 1125-1128 (1997).
11. H. C. Hang, C. Yu, M. R. Pratt, C. R. Bertozzi, Probing glycosyltransferase activities with the Staudinger ligation. *J Am Chem Soc* **126**, 6-7 (2004).
12. D. D. Park *et al.*, Membrane glycomics reveal heterogeneity and quantitative distribution of cell surface sialylation. *Chem Sci* **9**, 6271-6285 (2018).
13. C. Büll *et al.*, Targeting aberrant sialylation in cancer cells using a fluorinated sialic acid analog impairs adhesion, migration, and in vivo tumor growth. *Mol Cancer Ther* **12**, 1935-1946 (2013).
14. Q. Zhou, Y. Xie, M. Lam, C. B. Lebrilla, -Glycomic Analysis of the Cell Shows Specific Effects of Glycosyl Transferase Inhibitors. *Cells* **10**, (2021).
15. Y. Zhou *et al.*, Inhibition of fucosylation by 2-fluorofucose suppresses human liver cancer HepG2 cell proliferation and migration as well as tumor formation. *Sci Rep* **7**, 11563 (2017).
16. T. de Vries, R. M. Knegt, E. H. Holmes, B. A. Macher, Fucosyltransferases: structure/function studies. *Glycobiology* **11**, 119R-128R (2001).
17. J. Holgersson, A. Gustafsson, M. E. Breimer, Characteristics of protein-carbohydrate interactions as a basis for developing novel carbohydrate-based antirejection therapies. *Immunol Cell Biol* **83**, 694-708 (2005).
18. A. Varki, Biological roles of glycans. *Glycobiology* **27**, 3-49 (2017).
19. H. J. An *et al.*, Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol Cell Proteomics* **11**, M111.010660 (2012).

20. K. Zheng, C. Bantog, R. Bayer, The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs* **3**, 568-576 (2011).
21. A. M. Goetze *et al.*, High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* **21**, 949-959 (2011).
22. Z. Wang, J. Zhu, H. Lu, Antibody glycosylation: impact on antibody drug characteristics and quality control. *Appl Microbiol Biotechnol* **104**, 1905-1914 (2020).
23. T. T. Wang, IgG Fc Glycosylation in Human Immunity. *Curr Top Microbiol Immunol* **423**, 63-75 (2019).
24. R. Yogo *et al.*, The Fab portion of immunoglobulin G contributes to its binding to Fcγ receptor III. *Sci Rep* **9**, 11957 (2019).
25. D. Reusch, M. L. Tejada, Fc glycans of therapeutic antibodies as critical quality attributes. *Glycobiology* **25**, 1325-1334 (2015).
26. R. Jefferis, Isotype and glycoform selection for antibody therapeutics. *Arch Biochem Biophys* **526**, 159-166 (2012).
27. A. Beck, O. Cochet, T. Wurch, GlycoFi's technology to control the glycosylation of recombinant therapeutic proteins. *Expert Opin Drug Discov* **5**, 95-111 (2010).
28. A. Delobel, Glycosylation of Therapeutic Proteins: A Critical Quality Attribute. *Methods Mol Biol* **2271**, 1-21 (2021).
29. K. A. Karlsson, Pathogen-host protein-carbohydrate interactions as the basis of important infections. *Adv Exp Med Biol* **491**, 431-443 (2001).
30. K. Kato, A. Ishiwa, The role of carbohydrates in infection strategies of enteric pathogens. *Trop Med Health* **43**, 41-52 (2015).
31. Y. Rossez *et al.*, Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for *Helicobacter pylori*. *Glycobiology* **22**, 1193-1206 (2012).
32. A. K. Singh, S. H. Harrison, J. S. Schoeniger, Gangliosides as receptors for biological toxins: development of sensitive fluoroimmunoassays using ganglioside-bearing liposomes. *Anal Chem* **72**, 6019-6024 (2000).
33. J. Sjögren, M. Collin, Bacterial glycosidases in pathogenesis and glycoengineering. *Future Microbiol* **9**, 1039-1051 (2014).
34. A. Marcobal, A. M. Southwick, K. A. Earle, J. L. Sonnenburg, A refined palate: bacterial consumption of host glycans in the gut. *Glycobiology* **23**, 1038-1046 (2013).
35. P. Gagneux, A. Varki, Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9**, 747-755 (1999).
36. F. Broszeit *et al.*, Glycan remodeled erythrocytes facilitate antigenic characterization of recent A/H3N2 influenza viruses. *Nat Commun* **12**, 5449 (2021).
37. B. Ballester, J. Milara, J. Cortijo, The role of mucin 1 in respiratory diseases. *Eur Respir Rev* **30**, (2021).
38. Y. Kyo *et al.*, Anti-inflammatory role of MUC1 mucin during infection with nontypeable *Haemophilus influenzae*. *Am J Respir Cell Mol Biol* **46**, 149-156 (2012).
39. Y. Li, D. L. Dinwiddie, K. S. Harrod, Y. Jiang, K. C. Kim, Anti-inflammatory effect of MUC1 during respiratory syncytial virus infection of lung epithelial cells in vitro. *Am J Physiol Lung Cell Mol Physiol* **298**, L558-563 (2010).
40. W. Lu *et al.*, Cutting edge: enhanced pulmonary clearance of *Pseudomonas aeruginosa* by Muc1 knockout mice. *J Immunol* **176**, 3890-3894 (2006).

41. E. P. Lillehoj, K. Kato, W. Lu, K. C. Kim, Cellular and molecular biology of airway mucins. *Int Rev Cell Mol Biol* **303**, 139-202 (2013).
42. K. S. Bergstrom, L. Xia, Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology* **23**, 1026-1037 (2013).
43. E. P. Bennett *et al.*, Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22**, 736-756 (2012).
44. M. E. Johansson, H. Sjövall, G. C. Hansson, The gastrointestinal mucus system in health and disease. *Nat Rev Gastroenterol Hepatol* **10**, 352-361 (2013).
45. P. Palese, R. W. Compans, Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): mechanism of action. *J Gen Virol* **33**, 159-163 (1976).
46. S. Uchiyama *et al.*, The surface-anchored NanA protein promotes pneumococcal brain endothelial cell invasion. *J Exp Med* **206**, 1845-1852 (2009).
47. S. Vaishnava *et al.*, The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255-258 (2011).
48. L. G. Baum, O. B. Garner, K. Schaefer, B. Lee, Microbe-Host Interactions are Positively and Negatively Regulated by Galectin-Glycan Interactions. *Front Immunol* **5**, 284 (2014).
49. R. Medzhitov, C. A. Janeway, Innate immunity: the virtues of a nonclonal system of recognition. *Cell* **91**, 295-298 (1997).
50. C. A. Janeway, R. Medzhitov, Innate immune recognition. *Annu Rev Immunol* **20**, 197-216 (2002).
51. H. Kumar, T. Kawai, S. Akira, Toll-like receptors and innate immunity. *Biochem Biophys Res Commun* **388**, 621-625 (2009).
52. B. A. Beutler, TLRs and innate immunity. *Blood* **113**, 1399-1407 (2009).
53. G. Jengo, R. Bataille, A. Geffroy-Luseau, G. Descamps, C. Pellat-Deceunynck, Pathogen-associated molecular patterns are growth and survival factors for human myeloma cells through Toll-like receptors. *Leukemia* **20**, 1130-1137 (2006).
54. S. Akira, H. Hemmi, Recognition of pathogen-associated molecular patterns by TLR family. *Immunol Lett* **85**, 85-95 (2003).
55. B. K. Davis, H. Wen, J. P. Ting, The inflammasome NLRs in immunity, inflammation, and associated diseases. *Annu Rev Immunol* **29**, 707-735 (2011).
56. R. Caruso, N. Warner, N. Inohara, G. Núñez, NOD1 and NOD2: signaling, host defense, and inflammatory disease. *Immunity* **41**, 898-908 (2014).
57. O. Takeuchi, S. Akira, Pattern recognition receptors and inflammation. *Cell* **140**, 805-820 (2010).
58. A. Varki, Since there are PAMPs and DAMPs, there must be SAMPs? Glycan “self-associated molecular patterns” dampen innate immunity, but pathogens can mimic them. *Glycobiology* **21**, 1121-1124 (2011).
59. I. Secundino *et al.*, Host and pathogen hyaluronan signal through human siglec-9 to suppress neutrophil activation. *J Mol Med (Berl)* **94**, 219-233 (2016).
60. D. Bojar, R. K. Powers, D. M. Camacho, J. J. Collins, Deep-Learning Resources for Studying Glycan-Mediated Host-Microbe Interactions. *Cell Host Microbe* **29**, 132-144.e133 (2021).
61. S. Chen, R. Qin, L. K. Mahal, Sweet systems: technologies for glycomic analysis and their integration into systems biology. *Crit Rev Biochem Mol Biol* **56**, 301-320 (2021).

62. G. Lu, C. L. Crihfield, S. Gattu, L. M. Veltri, L. A. Holland, Capillary Electrophoresis Separations of Glycans. *Chem Rev* **118**, 7867-7885 (2018).
63. N. Nemanichvili *et al.*, Fluorescent Trimeric Hemagglutinins Reveal Multivalent Receptor Binding Properties. *J Mol Biol* **431**, 842-856 (2019).
64. K. T. Pilobello, D. E. Slawek, L. K. Mahal, A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. *Proc Natl Acad Sci U S A* **104**, 11534-11539 (2007).
65. K. T. Pilobello, L. Krishnamoorthy, D. Slawek, L. K. Mahal, Development of a lectin microarray for the rapid analysis of protein glycopatterns. *Chembiochem* **6**, 985-989 (2005).
66. A. Shajahan, C. Heiss, M. Ishihara, P. Azadi, Glycomic and glycoproteomic analysis of glycoproteins-a tutorial. *Anal Bioanal Chem* **409**, 4483-4505 (2017).
67. Q. Li, Y. Xie, M. Wong, M. Barboza, C. B. Lebrilla, Comprehensive structural glycomic characterization of the glycocalyxes of cells and tissues. *Nat Protoc* **15**, 2668-2704 (2020).
68. J. M. Suski *et al.*, Isolation of plasma membrane-associated membranes from rat liver. *Nat Protoc* **9**, 312-322 (2014).
69. Y. Fujiki, A. L. Hubbard, S. Fowler, P. B. Lazarow, Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J Cell Biol* **93**, 97-102 (1982).
70. Q. Li, Y. Xie, M. Wong, C. B. Lebrilla, Characterization of Cell Glycocalyx with Mass Spectrometry Methods. *Cells* **8**, (2019).
71. M. Uruse *et al.*, Phase separation of myelin sheath in Triton X-114 solution: predominant localization of the 21.5-kDa isoform of myelin basic protein in the lipid raft-associated domain. *J Biochem* **155**, 265-271 (2014).
72. C. Bordier, Phase separation of integral membrane proteins in Triton X-114 solution. *J Biol Chem* **256**, 1604-1607 (1981).
73. Y. Kim *et al.*, Use of colloidal silica-beads for the isolation of cell-surface proteins for mass spectrometry-based proteomics. *Methods Mol Biol* **748**, 227-241 (2011).
74. T. H. Plummer, J. H. Elder, S. Alexander, A. W. Phelan, A. L. Tarentino, Demonstration of peptide:N-glycosidase F activity in endo-beta-N-acetylglucosaminidase F preparations. *J Biol Chem* **259**, 10700-10704 (1984).
75. A. L. Tarentino, C. M. Gómez, T. H. Plummer, Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Biochemistry* **24**, 4665-4671 (1985).
76. V. Tretter, F. Altmann, L. März, Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1----3 to the asparagine-linked N-acetylglucosamine residue. *Eur J Biochem* **199**, 647-652 (1991).
77. W. Song *et al.*, N-glycan occupancy of Arabidopsis N-glycoproteins. *J Proteomics* **93**, 343-355 (2013).
78. T. Wang *et al.*, Discovery and characterization of a novel extremely acidic bacterial N-glycanase with combined advantages of PNGase F and A. *Biosci Rep* **34**, e00149 (2014).
79. S. Sun, J. Y. Zhou, W. Yang, H. Zhang, Inhibition of protein carbamylation in urea solution using ammonium-containing buffers. *Anal Biochem* **446**, 76-81 (2014).
80. Q. Li *et al.*, Site-Specific Glycosylation Quantitation of 50 Serum Glycoproteins Enhanced by Predictive Glycopeptidomics for Improved Disease Biomarker Discovery. *Anal Chem* **91**, 5433-5445 (2019).

81. Y. Du, F. Wang, K. May, W. Xu, H. Liu, LC-MS analysis of glycopeptides of recombinant monoclonal antibodies by a rapid digestion procedure. *J Chromatogr B Analyt Technol Biomed Life Sci* **907**, 87-93 (2012).
82. N. H. Packer, M. J. Harrison, Glycobiology and proteomics: is mass spectrometry the Holy Grail? *Electrophoresis* **19**, 1872-1882 (1998).
83. M. C. Hennion, Graphitized carbons for solid-phase extraction. *J Chromatogr A* **885**, 73-95 (2000).
84. L. R. Ruhaak, G. Xu, Q. Li, E. Goonatileke, C. B. Lebrilla, Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem Rev* **118**, 7886-7930 (2018).
85. E. Ruiz-May, C. Catalá, J. K. Rose, N-glycoprotein enrichment by lectin affinity chromatography. *Methods Mol Biol* **1072**, 633-643 (2014).
86. A. Bermudez, S. J. Pitteri, Enrichment of Intact Glycopeptides Using Strong Anion Exchange and Electrostatic Repulsion Hydrophilic Interaction Chromatography. *Methods Mol Biol* **2271**, 107-120 (2021).
87. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71 (1989).
88. T. Čaval, J. Zhu, A. J. R. Heck, Simply Extending the Mass Range in Electron Transfer Higher Energy Collisional Dissociation Increases Confidence in N-Glycopeptide Identification. *Anal Chem* **91**, 10401-10406 (2019).
89. Z. Chen *et al.*, Site-specific characterization and quantitation of N-glycopeptides in PKM2 knockout breast cancer cells using DiLeu isobaric tags enabled by electron-transfer/higher-energy collision dissociation (EThcD). *Analyst* **143**, 2508-2519 (2018).
90. K. L. Ford, W. Zeng, J. L. Heazlewood, A. Bacic, Characterization of protein N-glycosylation by tandem mass spectrometry using complementary fragmentation techniques. *Front Plant Sci* **6**, 674 (2015).
91. C. K. Frese *et al.*, Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal Chem* **84**, 9668-9673 (2012).
92. M. Bern, Y. J. Kil, C. Becker, Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinformatics* **Chapter 13**, Unit13.20 (2012).
93. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).
94. L. He, L. Xin, B. Shan, G. A. Lajoie, B. Ma, GlycoMaster DB: software to assist the automated identification of N-linked glycopeptides by tandem mass spectrometry. *J Proteome Res* **13**, 3881-3895 (2014).

Chapter 2 Host Cell Glycocalyx Remodeling Reveals SARS-Cov-2 Spike Protein Glycomic

Binding Sites

Authors:

Ying Sheng^{1, 2}; Anita Vinjamuri¹; Michael Russelle S. Alvarez³; Yixuan (Axe) Xie¹; Marisa McGrath⁴; Siyu (Cathy) Chen¹; Mariana Barboza^{1, 5}; Matthew Frieman⁴; Carlito B Lebrilla*^{1, 2}

¹Department of Chemistry, University of California, Davis, CA

²The Biochemistry, Molecular, Cellular and Developmental Biology (BMADB) Graduate Group, University of California, Davis, CA

³University of the Philippines Los Baños, Los Baños, Philippines

⁴Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD

⁵Department of Anatomy, Physiology & Cell Biology, School of Veterinary Medicine, University of California, Davis, CA

ABSTRACT

Glycans on the host cell membrane and viral proteins play critical roles in pathogenesis. Highly glycosylated epithelial cells represent the primary boundary separating embedded host tissues from pathogens within the respiratory and intestinal tracts. SARS-CoV-2, the causative agent for the COVID-19 pandemic, reaches into the respiratory tract. We found purified human milk oligosaccharides (HMOs) inhibited the viral binding on cells. Spike (S) protein receptor binding domain (RBD) binding to host cells were partly blocked by co-incubation with exogenous HMOs, most by 2-6-sialyl-lactose (6'SL), supporting the notion that HMOs can function as decoys

in defense against SARS-Cov2. To investigate the effect of host cell glycocalyx on viral adherence, we metabolically modified and confirmed with glycomic methods the cell surface glycome to enrich specific N-glycan types including those containing sialic acids, fucose, mannose, and terminal galactose. Additionally, Immunofluorescence studies demonstrated that the S protein preferentially binds to terminal sialic acids with α -(2,6)-linkages. Furthermore, site-specific glycosylation of S protein RBD and its human receptor ACE2 were characterized using LC-MS/MS. We then performed molecular dynamics calculations on the interaction complex to further explore the interactive complex between ACE2 and the S protein. The results showed that hydrogen bonds mediated the interactions between ACE2 glycans and S protein with desialylated glycans forming significantly fewer hydrogen bonds. These results supported a mechanism where the virus binds initially to glycans on host cells preferring α -(2,6)-sialic acids and finds ACE2 and with the proper orientation infects the cell.

INTRODUCTION

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), the causative agent of COVID-19(1), encodes an extensively glycosylated spike (S) protein that protrudes from the viral surface and binds angiotensin-converting enzyme 2 (ACE2) on host cells(2-6). This novel SARS-CoV-2 was found to share similarities with the SARS-CoV, which was responsible for the SARS pandemic that occurred in 2002(7, 8). ACE2 serves as the entry point for several coronaviruses into cells, including SARS-CoV and SARS-CoV-2(9, 10). The receptor binding domain (RBD) of SARS-CoV-2 S protein has been limited to amino acid residues Arg319 to Phe541(11-13). *In vitro* binding measurements also showed that the SARS-CoV-2 RBD binds to ACE2 with an affinity in the low nanomolar range, indicating that the RBD is a key functional component within the S1 subunit responsible for the binding of SARS-CoV-2 to ACE2(2, 13). The plasma membrane protein ACE2 is abundantly expressed in humans tissues, including respiratory and intestinal epithelia, liver arteries, heart and kidney(14).

Mammalian epithelial cells are highly glycosylated(15, 16) due to glycoproteins and glycolipids found on the cell membrane. Both the ACE2 receptor and the S protein are similarly extensively glycosylated. Several glycosylation sites are found near the binding interface(12, 17-19). The role of glycosylation in the interaction between human ACE2 and SARS-CoV-2 S protein has been extensively studied, primarily using molecular dynamics (MD) simulations(12, 20, 21). Human ACE2 variants have also been modeled, characterized, and examined for susceptibility to coronavirus interactions(22, 23). Among ACE2 glycosylation sites, one of the most characterized position for its role in S protein binding and viral infectivity is the asparagine on position 90 (N90). Recent genetic and biochemical studies showed that mutations that removed glycosylation on

N90 site directly increased the susceptibility to SARS-CoV-2 infection(21, 23). In contrast, glycans present on N322 and N90 have the opposite effects on S protein binding. The N322 glycan interacts tightly with the RBD of the ACE2-bound S protein and strengthens the complex(20). The S protein also contains glycosaminoglycan (GAG) binding motifs so that host surface GAGs contribute to cell entry by SARS-CoV-2 (24). Additionally, heparan sulfate has also been shown to promote Spike-ACE2 interaction (25).

Pathogen adhesion is often mediated by highly specific lectin-glycan interactions. For example, *Escherichia coli* with type 1 fimbriae binds to cell surfaces exhibiting preference for high mannose glycans, while *Escherichia coli* with type S fimbriae has binding specificity for α -(2,3)-linked sialic acids. Cell surface glycans have also been shown to act as a shield to mask its identity as a viable host to the pathogen. It was recently proposed that HMOs can prevent viral adhesion to intestinal epithelial cells via binding to the epithelial surface, causing structural changes in the receptor thereby impeding the virus from hijacking the host cell(26). Breast-fed infants have significant amounts of HMOs lining the mucosal surface of their gastrointestinal tract. While the viral binding to glycans and HMO in particular have been studied, the direct interaction between the virus and host glycans remain relatively unexplored.

In this study, the role of host glycosylation and its effect on S protein binding was examined by identifying the host glycans that are involved in the binding. The study began with HMOs in a rapid assay to determine the broad details of the oligosaccharide that bind the virus. We then examined the impact of host cell glycosylation on S protein binding, by modifying the host glycosylation while leaving protein expression unchanged using transferase inhibitors. Using newly developed methods glycomic tools, we found that specific glycans on the host cell facilitate

S protein binding and that binding depends more on the nature of glycans than it does on the membrane proteins.

METHODS AND MATERIALS

HMO Purification

HMOs were obtained from breast milk samples using previously reported methods(27, 28). Briefly, breast milk samples from 7 mothers were pooled. Pooled sample was defatted through centrifugation, proteins were precipitated with ethanol, and the resulting glycans were reduced with sodium borohydride (Sigma-Aldrich, St. Louis, MO). Solid phase extraction was performed on 25 mg graphitized carbon cartridges (ThermoFisher). Solvents were dried in vacuo using miVac (SP Scientific, PA) and purified HMOs were reconstituted and diluted prior to analysis.

Inhibition of HMO against SARS-CoV-2

All HMO screens were performed with Vero E6 cells. Cells were plated in 96 well plates at 5×10^3 cells/well one day prior to infection. HMOs were diluted from stock to 50 μ M and an 8-point 1:2 dilution series was prepared in duplicate in Vero Media. Every compound dilution and control were normalized to contain the same concentration of drug vehicle (e.g., DMSO). Cell plates were pre-treated with drug for 2 h at 37°C (5% CO₂) prior to infection with diluted SARS-CoV-2 GFP for a final MOI of 0.1. In addition to plates that were infected, parallel plates were left uninfected to monitor cytotoxicity of drug alone, measured by CellTiter-Glo (CTG) assays as per the manufacturer's instructions (Promega, Madison, WI). Plates were then incubated at 37°C (5% CO₂) for 48 hours, followed by fixation with 4.0% paraformaldehyde, nuclear staining with Hoechst (Invitrogen, Carlsbad, CA), and data acquisition on a Celigo 5-channel Imaging Cytometer (Nexcelom Bioscience, Lawrence, MA). The percent of infected cells was determined for each

well based on GFP expression by manual gating using the Celigo software. For the CTG assays, luminescence was read on a BioTek Synergy HTX plate reader (BioTek Instruments Inc., Winooski, VT) using the Gen5 software (v7.07, Biotek Instruments Inc., Winooski, VT).

Cell culture and glycocalyx remodeling treatments

Human liver hepatocellular carcinoma HepG2, lung carcinoma epithelial Calu-3, urinary bladder epithelial RT4 cells were obtained from American Type Culture Collection (ATCC, VA). HepG2 and Calu-3 cells were grown in Eagle's Minimum Essential Medium (EMEM). RT4 cells were cultured in McCoy's 5a Medium. All media were supplemented with 10% (v/v) fetal bovine serum and 100 U mL⁻¹ penicillin and streptomycin. Cells were subcultured at 90% confluency and maintained at 37 °C in a humidified incubator with 5% CO₂. At 50% cell confluency, the cells were either treated with 150 μM kifunensine, 2-fluoro-L-fucose, or 3-fluorinated sialic acid for 48 hours.

Sample Information

Recombinant human angiotensin-converting enzyme 2(ACE2), SARS-CoV-2 Spike protein S1 Subunit RBD (Arg319-Phe541) and Spike protein S1 subunit (Val16-Arg685) derived from transfected human HEK293 cells were obtained from RayBiotech (Georgia, Product Number 230-30165, 230-30162) and Sino Biological (China, Product Number 40591-V08H), respectively.

Immunofluorescence

The cells were seeded into FluoroDish™ cell culture dishes (WPI, FL) coated with poly-d-lysine with appropriate density using EMEM cell culture media. At 40% confluency, cells were treated with media either supplemented with 150 μM kifunensine, 2-fluoro-L-fucose, or 3-

fluorinated sialic acid for 48 hours. Control cell culture without treatment and treated cells were rinsed with phosphate-buffered saline (PBS), and fixed with 4% paraformaldehyde (Affymetrix, OH). Recombinant SARS-CoV-2 spike protein RBD and S1 subunits were conjugated to a fluorescent label with Alexa Fluor™ 555 according to manufacturing instructions (Microscale Protein Labeling Kit, Invitrogen, MA). Fixed control and glyco-modified cells were then incubated with fluorescent labelled S proteins or Anti-ACE2 antibody (Santa Cruz Biotechnology, TX) in PBS at 4°C for 18 hours. Cells were stained for the nucleus with 1.6 µM Hoechst 33342 (Thermo Fisher Scientific, MA) followed by the staining for the plasma membrane with 1000-fold diluted CellMask™ Deep Red Plasma Membrane Stain (Thermo Fisher Scientific, MA), respectively at 37 °C for 10 min. Fluorescence images were captured using a Leica TCS SP8 STED 3X Super-Resolution Confocal Microscope (Wetzlar, Germany). Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ.

Cell membrane extraction

Cell membrane fractions were prepared as previously described (16, 29, 30). Briefly, control and glycoengineered cells were collected and resuspended in homogenization buffer containing 0.25 M sucrose, 20 mM HEPES-KOH (pH 7.4), and protease inhibitor mixture (1:100; Calbiochem/EMD Chemicals). Cells were lysed on ice with five alternating on and off pulses in 5 and 10 second intervals using a probe sonicator (Qsonica, CT). Nuclear and mitochondrial fractions and cellular debris were pelleted and isolated by centrifugation at 2000 × *g* for 10 min. The supernatants were then ultra-centrifuged at 200 000 × *g* for 45 min at 4 °C to extract the plasma membrane. The pellets of the cell membrane were resuspended in 500 µL of 0.2 M

Na₂CO₃ solution and 500 μL of water followed by two more ultracentrifugation treatments at 200 000 × *g* for 45 min to wash off the endoplasmic reticulum (ER) and cytoplasmic fraction.

Enzymatic N-glycan release and purification of N-glycans

Extracted cell membrane fractions or RNase B were suspended with 100 μL of 100 mM NH₄HCO₃ in 5 mM dithiothreitol and heated in boiling water for 2 minutes to denature the proteins. Solutions of with 2 μL of peptide N-glycosidase F (New England Biolabs, MA) were added to the samples to release the N-glycans from proteins, and the resulting solutions were then incubated in a microwave reactor (CEM Corporation, NC) at 20 watts, 37 °C for 10 min. The samples were further placed in a 37 °C water bath for 18 hours. Ultracentrifugation at 200 000 × *g* for 30 min was performed to precipitate membrane fractions, and the supernatant containing N-glycans was collected and purified using porous graphitic carbon (PGC) on a 96-well SPE plate (Grace, IL). The plate was equilibrated with 80% (v/v) acetonitrile containing 0.1% (v/v) trifluoroacetic acid. The samples were loaded onto the plate and washed with nanopure water. N-Glycans were eluted with a solution of 40% (v/v) acetonitrile containing 0.05% (v/v) trifluoroacetic acid, and the samples were dried in vacuo using miVac (SP Scientific, PA) prior to mass spectrometric analysis.

Glycoprotein digestion and enrichment

Details of the protein digestion have been described previously⁽³⁰⁾. Extracted cell membrane proteins were reconstituted in 60 μL of 8 M urea at room temperature. Recombinant proteins and dissolved cell membrane proteins were reduced with 2 μL of 550 mM dithiothreitol, alkylated with 4 μL of 450 mM iodoacetamide. A 420 μL of 50 mM ammonium bicarbonate

solution was added to dilute the urea concentration to 1 M and to adjust the pH value. The samples were incubated with 2 µg trypsin at 37 °C for 18 hours. The resulting peptides were concentrated *in vacuo* using miVac (SP Scientific, PA). Glycopeptides were enriched by solid-phase extraction using iSPE®-HILIC cartridges (HILICON, Sweden). The cartridges were conditioned with 0.1% (v/v) trifluoroacetic acid in acetonitrile, followed by 1% (v/v) trifluoroacetic acid and 80% (v/v) acetonitrile in water. The samples were loaded and washed with 1% (v/v) trifluoroacetic acid and 80% (v/v) acetonitrile in water. The enriched glycopeptides were eluted with water containing 0.1% (v/v) trifluoroacetic acid and dried prior to mass spectrometric analysis.

Glycomic analysis with LC-MS/MS

Glycan samples were reconstituted with 30 µL nanopure water and analyzed using an Agilent 6520 Accurate Mass Q-TOF LC/MS equipped with a PGC nano-chip (Agilent Technologies, CA). The glycan separation was performed at a constant flow rate of 300 nL min⁻¹, and a binary gradient was applied using (A) 0.1% (v/v) formic acid in 3% acetonitrile and (B) 1% (v/v) formic acid in 90% acetonitrile: 0–2 min, 0–0% (B); 2–20 min, 0–16% (B); 20–40 min, 16%–72% (B); 40–42 min, 72–100% (B); 42–52 min, 100–100% (B); 52–54 min, 100–0% (B); 54–65 min, 0–0% (B). MS spectra were collected with a mass range of m/z 600–2000 at a rate of 1.5 s per spectrum in positive ionization mode. The most abundant precursor ions in each MS1 spectrum were subjected to fragmentation through collision-induced dissociation (CID) based on the equation $V_{\text{collision}} = 1.8 \times (m/z) / 100 V - 2.4 V$.

Glycomic data analysis

Extraction of the compound chromatographs of glycans from cells was obtained via the MassHunter Qualitative Analysis B08 software (Agilent, CA). N-Glycan compositions were identified according to accurate masses using an in-house library constructed based on the knowledge of N-glycan biosynthetic pathways and previously obtained in-house structures of N-glycans. Relative abundances were determined by integrating peak areas for observed glycan masses and normalizing to the summed peak areas of all glycans detected.

Glycoproteomic analysis with LC-MS/MS

The enriched glycopeptide samples were reconstituted with nanopure water and directly characterized using UltiMate™ WPS-3000RS nanoLC 980 system coupled to the Nanospray Flex ion source of an Orbitrap Fusion Lumos Tribrid Mass Spectrometer system (Thermo Fisher Scientific, MA). The analytes were separated on an Acclaim™ PepMap™ 100 C18 LC Column (3 μm, 0.075 mm x 150 mm, ThermoFisher Scientific). A binary gradient was applied using 0.1% (v/v) formic acid in (A) water and (B) 80% acetonitrile: 0–5 min, 4–4% (B); 5–133 min, 4–32% (B); 133–152 min, 32%–48% (B); 152–155 min, 48–100% (B); 155–170 min, 100–100% (B); 170–171 min, 100–4% (B); 171–180 min, 4–4% (B). The instrument was run in data-dependent mode with 1.8kV spray voltage, 275 °C ion transfer capillary temperature, and the acquisition was performed with the full MS scanned from 700 to 2000 in positive ionization mode. Stepped higher-energy C-trap dissociation (HCD) at 30±10% was applied to obtain tandem MS/MS spectra with m/z values starting from 120.

Glycoproteomic data analysis

Glycopeptide fragmentation spectra were annotated using Byonic software (Protein Metrics, CA) against the reviewed UniProt Severe acute respiratory syndrome coronavirus 2 spike protein database. Carbamidomethyl modification at cysteine residues and oxidation at methionine were assigned as the modification.

Molecular dynamic simulation of S protein on ACE2

The 3D structure of S protein and ACE2 complex was obtained from PDB (PDB code 7DF4) (31). The most abundant glycans for each ACE2 glycosite were modeled and attached to the protein using CHARMM-GUI(32). Additionally, the fully-desialylated glycans were modeled and attached to generate a fully-desialylated homolog of the ACE2 glycoprotein. The models were solvated using the TIP3P water model, and counterions were added to neutralize the system. The CHARMM carbohydrate force field(33) and CHARMM36m force field(34) were used for the carbohydrate and protein structures. Equilibration was performed at 303.15 K over 10 ps. Molecular dynamics simulation was performed using NAMD software package version 2.13(35) at 303.15 K under NPT conditions over 5 ns with an output every 10 ps. Long-range electrostatics were evaluated using the particle-mesh Ewald (PME) method(36). Covalent bonds involving hydrogen were constrained with the SHAKE algorithm(37). After dynamics simulations, trajectories were loaded onto VMD for visualization and analysis(38). Specifically, the intermolecular hydrogen-bonding interactions (donor-acceptor distance 3.0 Å, angle cutoff 20°) of each glycan in the fully-sialylated and desialylated forms were compared over the simulation period.

RESULTS

Inhibition of virus binding by human milk oligosaccharides

Human milk oligosaccharides (HMOs) contain a number of unique structures that can be used to rapidly screen the glycan specificity of the virus adhesion. We tested whether the SARS-CoV-2 virus could be inhibited by HMOs. We first examined whether pooled samples of purified HMOs from seven different mothers could affect the binding of SARS-CoV-2 virus on Vero E6 cells. **Figure 1a** showed that the binding capability was lower caused by the HMO mixture to about 25%. HMOs contain compounds with terminal fucose, sialic acid and galactose. To identify the functional components that could specifically affect binding, we further tested individual compounds that contained these terminal saccharides. The HMOs 2'-fucosyllactose (2'-FL), 6'-sialyllactose (6'-SL), and lacto-N-neotetraose (LNnT) were selected for this study because they represent many of the structures and are abundant in mothers' milk. 2'-FL and 6'-SL were produced by adding fucose or a N-acetylneuraminic acid (Neu5Ac)(39) to the lactose core, respectively. Lacto-N-neotetraose (LNnT) is a neutral HMO with a galactose terminus and contained neither fucose nor sialic acid. The infection studies showed that 2'-FL did not diminish infection, however both 6'-SL and LNnT showed some diminished infection to a similar extent as the pooled sample (**Figure 2.1**).

Due to limitations with working on the whole intact virus, we moved the research towards using the S protein as a surrogate for the virus. To validate this model, we performed the experiments on the S protein using the fluorescent labeling and immunofluorescence imaging. SARS-CoV-2 enters host cells via the angiotensin-converting enzyme 2 (ACE2) receptor, which binds the receptor binding domain (RBD) of the S protein(12). The Human Protein Atlas (HPA), a

website resource for protein expression profiles in cells, tissues and organs (<https://www.proteinatlas.org/>)(40, 41) was used to select the host cell with ACE2 expression. HepG2 was selected after confirming ACE2 expression with labeled antibody and immunofluorescence on the cell membrane (**Figure S2.1**).

In order to verify further whether HMOs block viral adhesion, we tested the ability of the selected HMO compounds to inhibit RBD binding to HepG2 cells with immunofluorescence. Preincubating HepG2 cells with HMOs did not decrease the binding between the RBD and the cells suggesting that the HMOs did not block binding sites on the host cell surface (**Figure S2.2**). We then tested whether the HMOs could block or alter the RBD of the virus by preincubating the RBD and the HMOs before introduction to HepG2 cells. Fluorescently labelled RBD was preincubated with 2'-FL, LNnT and 6'-SL separately then allowed to interact with host cells (**Figure 2.2A**). Quantitation of fluorescent signal intensity showed that HMOs blocked binding of RBD to cells presumably reflecting the behavior of the intact virus. The RBD was blocked only slightly by 2'FL (not statistically significant), more by LNnT (significant), and the most by 6'SL (**Figure 2.2B**). The data further showed that HMOs can potentially function as decoys to affect SARS-CoV-2 adherence.

Determining SARS-CoV-2 binding through variable glycoalyx expression

The notable decrease in binding caused by 6'-SL drew our attention to sialic acids as potential receptors on the cell surface. To further investigate the effects of cell surface glycans on RBD binding, we altered the cell membrane glycans through transferase inhibitors. We first characterized the glycan of the cell membrane and ACE2 on the native cell line. For this analysis, complex and hybrid type glycans were combined to distinguish them from oligomannose type.

The N-glycan profile shows a notable abundance of sialylated and sialyfucoylated structures (**Figure 3A**). The most abundant N-glycan compositions had multiple fucose and sialic acid (N-acetylneuraminic acid or Neu5Ac) residues such as Hex₆HexNAc₅Fuc₂NeuAc₃, Hex₆HexNAc₅Fuc₁NeuAc₃ and Hex₅HexNAc₄Fuc₁NeuAc₂. Glycoproteomic analysis of the cell membrane revealed seven glycosites on the ACE2 protein of HepG2 cells. The N-glycoforms of the ACE2 protein extracted from HepG2 cells were diverse and the most common structures were both fucosylated and sialylated (**Figure 2.3B, Table S2.1**). For comparison, we analyzed the glycosylation of commercial recombinant ACE2 protein expressed from HEK293 (**Figure S2.2, Table S2.2**) and found them to be similar to those expressed by HepG2 (**Table 2.1**). Both proteins were highly sialylated and fucosylated with limited amounts of high-mannose glycans.

We metabolically altered the cell surface glycome to enrich for sialic acids, fucoses and mannoses, respectively. To determine whether these changes in glycosylation affected ACE2 expression on the cell membrane, we probed the cells with fluorescently labeled antibodies (**Figure S2.4**). These experiments showed no significant changes in protein expression for ACE2 in any of the glycan modification procedures. To diminish fucosylation on the HepG2 cell surface, we employed a fucosyltransferase inhibitor, 2-fluoro-L-fucose (2F-Fucose). To inhibit sialylation, a sialyltransferase inhibitor 3-fluorinated sialic acid(3-F-Sia) was used.

The predicted behavior of each substrate are shown in **Figure 2.4a**. Compositional profiles were generated for the modified cells, using the sum of the intensities for similar glycan types from the LC-MS analysis. These inhibitors have recently been applied for altering cell surface glycosylation to yield similar results (42) (**Figure 2.4B**). 2F-Fucose inhibits fucosylation by being converted to the sugar nucleotide GDP-2F-Fuc(43). It then accumulates in the cell and binds to

the transferase and prohibits the enzyme from adding fucose to the nascent chain, thereby decreasing fucose expression on the cell surface (44). The sialyfucoylated N-glycans decreased from 75% to 10% after inhibition with 2F-Fucose treatment. The sialyfucoylated N-glycans were converted to sialylated (only) ones. For example, the abundant sialyfucoylated compound Hex₅HexNAc₄Fuc₁NeuAc decreased (9.6% to 1.9%, relative abundance) relative to the unfucoylated species Hex₅HexNAc₄Fuc₀NeuAc₂ which increased in abundance from 3.7% to 19% (**Table S2.3**). The sialylation pathway was inhibited using 3-F-Sia, a fluorinated sialic acid substrate [cytidine monophosphate (CMP)–SiaFAc](45), which binds more strongly to the enzyme thereby prohibiting the transfer of sialic acids. Treatment with 3-F-Sia decreased the relative abundances of all sialyfucoylated N-glycans from 75% to 34%. Simultaneously, the relative abundance of fucoylated (only) species increased from 1 % to 27 %. Thus, it appears that the inhibitors are highly effective diminishing fucoylated and sialylated structures, respectively.

After confirming that glycan alteration had taken place in host cell, immunofluorescence analysis was used to observe the effect of host glycome on viral binding. Treatment of 2F-Fucose did not affect RBD binding to the cell significantly as observed by immunofluorescence imaging (**Figure 2.5A**). However, inhibition of sialylation by 3-F-Sia decreased the S protein RBD binding with HepG2 cells by 64% (**Figure 2.5B**), indicating that the binding was likely mediated by sialic acid residues on the host cell surface. Similar trends were observed in other cell lines with ACE2 expression, namely Calu3 and RT4. Desialylation inhibited the binding from S protein RBD significantly, and decreased fucoylation did not change the extent of the binding.

In mammalian cells, terminal sialic acids are commonly found in α -glycosidic linkage to the C-3 or C-6 hydroxyl of galactose via α -(2,3)- or α -(2,6)-linkage for N-glycans (46). In nasal

mucosa, α -(2-6)-sialic acids are dominant(47) with significantly less detected in the lung(48). We further investigated linkage specificities for RBD binding by preincubating the RBD with sialylated HMOs. Preincubation with 3'-sialyllactose (3'-SL) did not decrease the binding, while significant decrease was observed after preincubation with 6'-SL, confirming S protein RBD prefers binding with α -(2-6) sialic acids (**Figure 2.6**).

Fucosylated glycans were also observed on ACE2 proteins in HepG2 cells (**Table S2.1**). Terminal α -(1,2) and α -(1,3)-fucose residues are commonly found in mammalian cells(49, 50). To confirm that fucosylation is less important, 2'-FL and 3'-fucosyllactose (3'-FL), components of HMOs, were used(39). Preincubating the S protein RBD with 2'-FL or 3'-FL did not significantly alter binding (**Figure S2.5**). The S protein RBD again showed little affinity to terminal fucose residues on host cells.

Glycosylation of HepG2 included primarily complex and hybrid type structures with fewer high mannose structures. The latter have been reported as important mediators in host-virus binding for human coronaviruses HKU1(51) and severe acute respiratory syndrome (SARS)(52). We remodeled the cell surface to produce primarily oligomannose and determined its effects on SARS-CoV-2 binding. Kifunensine (Kif) is commonly used to inhibit the α -mannosidase-I(53), thereby preventing mannose trimming to increase oligomannose-type glycans(52, 54). Our LC/MS data also proved its increasing the relative abundance of oligomannose to 89% in whole cell N-glycome as shown in **Figure 2.4**. Introduction of Kif to the cell resulted in a fourfold increase in the binding as measured by immunofluorescence imaging (**Figure 2.7A, 2.7B**). N-Glycans, released from RNase B, was also employed to examine high mannose type binding. Preincubation with the oligomannose decreased the binding of S protein RBD with host cells (**Figure 2.7C**). This

effect was dose dependent with higher concentrations preventing binding more strongly. High mannose glycans on host cell surface can therefore increase the adherence of S protein RBD.

To further validate the binding of the spike protein with the host glycoalyx, we used spike protein S1 subunit, a longer polypeptide segment of the S protein and includes the sequence of RBD. Treatment of the cell line with 2F-Fucose did not change the binding between S1 subunit and host cells (**Figure 2.7D**). Similarly, the use of 3-F-Sia significant decreased the fluorescent intensity of the assay demonstrating again that the spike protein binds to sialic acids. Surprisingly, the use of Kif on the S1 subunit no longer increased binding. The binding studies showed that there was no significant change in binding relative to the control.

Molecular Dynamics Calculations of ACE2 and S Protein Interactions

To gain further insight into the interactions between the primary receptor ACE2 (12, 55, 56) and the SARS-CoV-2 S protein, we performed molecular dynamics calculations on the interacting complex. Based on the glycoproteomic results for ACE2 from the HepG2 cell line (**Table S2.1**), we constructed a model with selected glycoforms on ACE2. ACE2 contained seven occupied N-glycan sites corresponding to Asn 53, 90, 103, 322, 432, 546, and 690 (**Figure 2.8A**). From the quantitative glycoproteomic results and the Protein Data Bank-derived complex (PDB ID: 7DF4)(31), the most abundant glycan at each site were modelled with CHARMM-GUI(32). The resulting structure, shown in the “up” conformation, was selected because it represented the activated complex prior to invasion. Molecular dynamics simulations were performed on the complex with solvent and associated ions for 5 ns (See Methods Section). After the simulations, the number of hydrogen bonds formed between the ACE2 glycans and S protein determined. For comparison, the same calculations were performed on the fully desialylated ACE2 homologs

(Figure 2.8B). The results showed that many of the glycans on ACE2 interacted with the S protein through hydrogen bonding interactions. Comparison of the fully sialylated and desialylated glycans showed significantly lower number of hydrogen bonds (based on 3 Å, donor-acceptor distances) particularly on Asn 90 (22 hydrogen bonds by glycan) and Asn 322 (51 hydrogen bonds by glycan) of the desialylated homolog **(Figure 2.9A)**. These results are consistent with earlier simulations performed by Zhao et al on ACE2 - S who noted that both glycan sites were also the most interactive in the complex(21). Furthermore, when the sialic acids were considered relative to other monosaccharide residues (3 by sialic acid at Asn 90, 15 by sialic acid at Asn 322), their contributions to the overall interactions were proportionally larger **(Figure 2.9B)**.

DISCUSSION

Glycans on the host cell membrane and on viral proteins play key roles in the infection of SARS-CoV-2. Viral glycosylation has been the primary focus of glycomic studies related to the virus. Indeed, the virus is highly glycosylated with at least 17 N-glycosylation and 2 O-glycosylation sites identified(19). We found two occupied N-glycosites on Spike RBD (**Table S2.4**) consistent with earlier findings. However, the host cells were also highly glycosylated. The LC-MS glycomic profile of HepG2 shows cell membrane with an abundance of high mannose-type glycans as well as complex-type structures with a high degree of sialylation. These structures are also branched with a combination of bi, tri, and higher antennary structures. The HepG2 cell lines was selected for its expression of ACE2, and these highly sialylated branched structures were similarly present in the protein further alluding to the importance of sialylation in at least the host-virus adhesion process.

The results showed that sialic acid in human milk oligosaccharides (HMOs) can block the binding of virus on the cell membrane. These results are further supported by recent findings that show similar deflecting properties of sialylated HMOs toward the S protein of SARS-CoV-2(26) and illustrating further the protective nature of human milk against these pathogens. HMOs are more similar to O-glycans in structure, however N-glycans on membrane proteins similarly provide sialic acid on their termini. Altering the glycans on the cell membrane, while maintaining the expression levels of proteins such as ACE2, further shows that sialic acid on the cell surface induces stronger binding to the virus. ACE2 is itself highly sialylated, in the cell line used in this study and from commercial sources (mainly from HEK293). ACE2 expressed recombinantly in other cell lines have similar glycosylation profiles that are similarly rich in sialylation. Deeper

structural analysis showed that the binding prefers a specific linkage, namely α -(2,6)-sialic acids. Interestingly, the human influenza virus has a similar preference for binding(57-59). Perhaps not coincidentally, the human epithelial is rich in α -(2,6)-sialic acid, which is also more abundant than the isomer α -(2,3)-sialic acid, the binding site of avian bird flu.

The binding of sialic acid point to specific protective measures by the host. In breast fed infants, HMOs can provide some protection. Human milk is also full of proteins that are highly sialylated such as the immunoglobulins and lactoferrins(60-63). In adults, pathogen deflection is performed by the mucus layer. SARS-CoV-2 is a respiratory disease reaching deep into the respiratory tract and the lungs. It also infects the intestine (64), with both types of tissues protected by a mucus layer constructed around high molecular weight glycoproteins called mucins(65). Mucin are expressed in epithelial surfaces of gastrointestinal, genitourinary, and respiratory tracts, where they also shield the surface against chemical and physical damages(46). While mucins are covered primarily by O-glycans that are similar to HMOs, they contain the same sialic acid termini as N-glycans. The mucus layer therefore presents a myriad of potential binding sites for commensal and pathogenic microbes(66, 67), and shedding mucins is a defense strategy against pathogen infection.

The high mannose glycans were also strongly bound in the shorter version (RBD) of the S protein. However, in the longer homolog (S1 subunit) this binding was diminished. These results suggest that that there is a high mannose binding site on the S protein that is potentially shielded in the longer homolog. On the other hand, high mannose glycans are typically not found on epithelial cells(16) and are not abundant in the blood. However, they are much more abundant in the tissue samples compared to serum. These glycans are found in cancer cells (68, 69) and

stem cells(70). The levels of several oligomannose type glycans are upregulated in tumor tissue(71, 72). The role of mannose residues as a host receptor has been demonstrated in the various microbe-host interactions, such as *Salmonella enterica* subsp. *enterica* serovar typhimurium (*S. typhimurium*)(73), influenza virus(74, 75), dengue virus(76) and human immunodeficiency virus (HIV)(77). Mannan is usually employed for studying mannose binding with virus(51, 52, 75, 78). The mannans are highly heterogeneous in length and branching. The repeating α -(1,6)-linked mannose backbone is usually branched by short chains of α -(1-2) and α -(1-3)-linked mannose structures(79). In this study, we used oligomannose released from RNase B(80) instead of mannan. The released high mannose glycans were determined with mass spectrometry (**Figure S2.6**), and all those structures have been found in human cell glycomes.

The integrated method developed here, which includes alteration of cell surface glycan products through specific inhibitors, coupled with the enrichment of the membrane proteins and extensive glycomic and glycoproteomic analysis provides a new platform for obtaining structural specificity in host-microbe interactions. Glycans are common targets for many commensals and pathogens alike. This method will have great utility in identifying glycan targets of individual microbes and even toxins that bind glycans. The method is made possible by recent advancements in novel glycosyl transferase inhibitors that produce specifically glycosylated membrane proteins. We noted that the conversion to a glycan type is never fully complete. There are residual endogenous glycans due to the differences in turnover of different glycoproteoform(81). However, the ability to perform glycomic profiling with LC-MS provides a guiding assay to examine the extent of the glycomic transformation.

CONCLUSIONS

The study supports a mechanism for binding of SARS-CoV-2 to the cell membrane that is primarily mediated by glycans. The preferred target of the S protein is sialylated glycans with α -(2,6)-sialic acids on the termini positions. The virus likely binds to cells and tissues rich in sialylated glycans, whether N-, O-, and potentially even glycolipids that are found in the surface of the epithelial surface. The airway epithelial is rich in sialic acids and in particularly α -(2,6)-sialic acids. In this regard, the human influenza virus and SARS-CoV-2 have the same binding preference in the host membranes. Invasion of SARS-CoV-2 likely occurs when the virus fortuitously binds to the ACE2 protein, which itself is highly sialylated. The alignment between the S and the ACE2 protein is further facilitated by hydrogen binding interactions between the sialylated glycans of the host cell and the polypeptide of the S protein.

ACKNOWLEDGMENT

The molecular dynamics calculations were carried out using the High-Performance Computing Cluster, COARE Facility of the Department of Science and Technology – Advance Science and Technology Institute, Philippines.

Funding: This work was supported by grants from the NIH (R01GM049077 to C.L.)

AUTHOR CONTRIBUTION

Ying Sheng wrote most of the manuscript. Anita Vinjamuri performed human milk oligosaccharides (HMOs) purification and contributed to manuscript editing. Michael Russelle S. Alvarez performed molecular dynamic simulation and contributed to the writing with his expertise. Yixuan Xie conceived the idea of exploring cell glycome affecting SARS-CoV2 binding and contributed to the glycomic data analysis. Marisa McGrath and Matthew Frieman contributed to the assay of HMO inhibiting against SARS-CoV-2. Siyu Chen contributed to purifying oligomannose compounds. Mariana Barboza conceived the idea of identifying the binding preference for sialic acid linkages and edited the manuscript according to her experience. Carlito B Lebrilla contributed to funding acquisition, writing—review and editing.

FIGURES

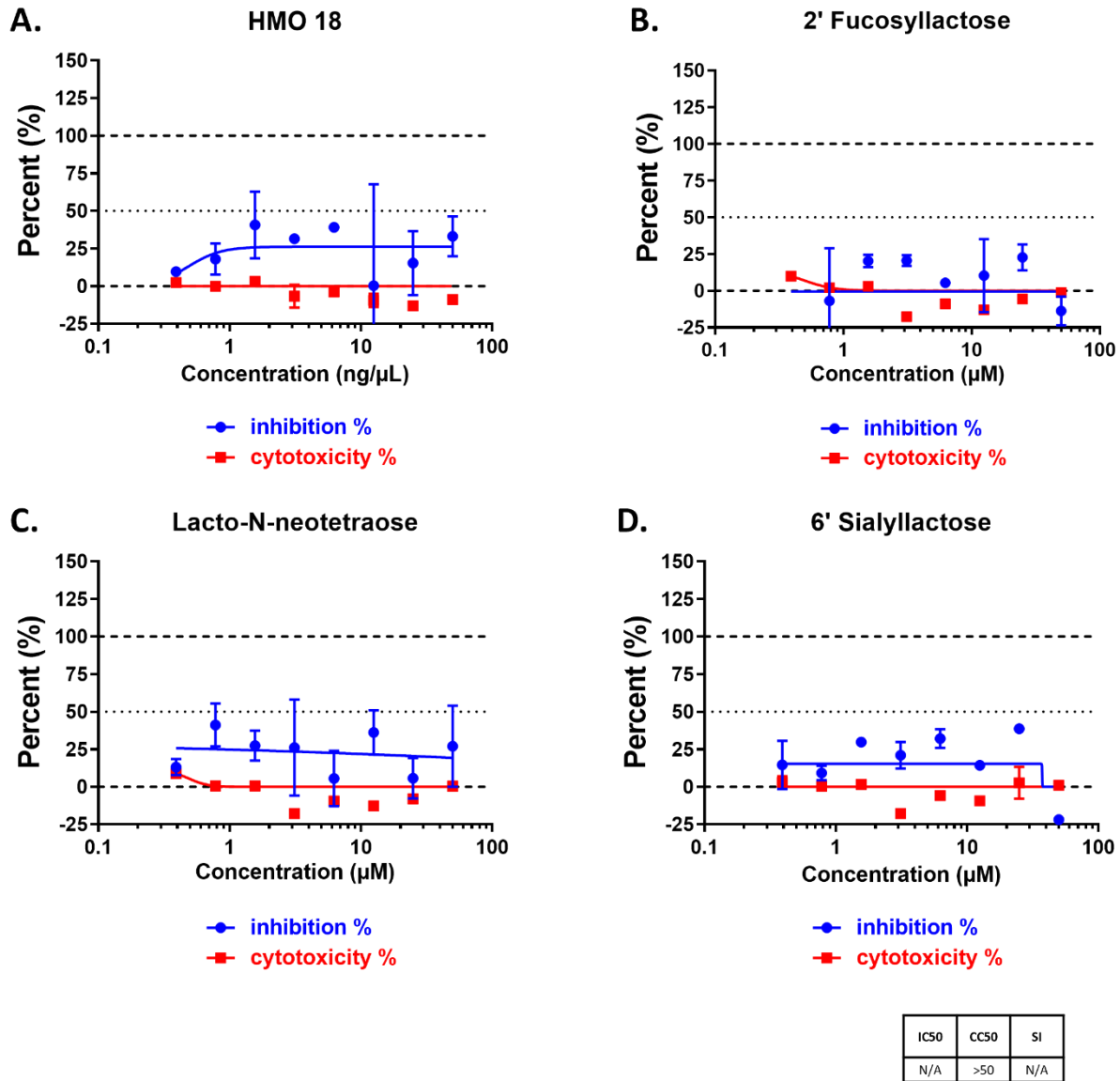


Figure 2.1. Viral infection on cells and cytotoxicity assays. Cell plates were pre-treated with pooled HMOs(A), 2'-FL(B), LNnT(C), and 6'-SL(C), respectively. The treatment was performed for 2 h at 37°C (5% CO₂) prior to infection. The percent of infected cells was determined for each well based on GFP expression. All samples were run in triplicate on both an assay plate and a toxicity plate.

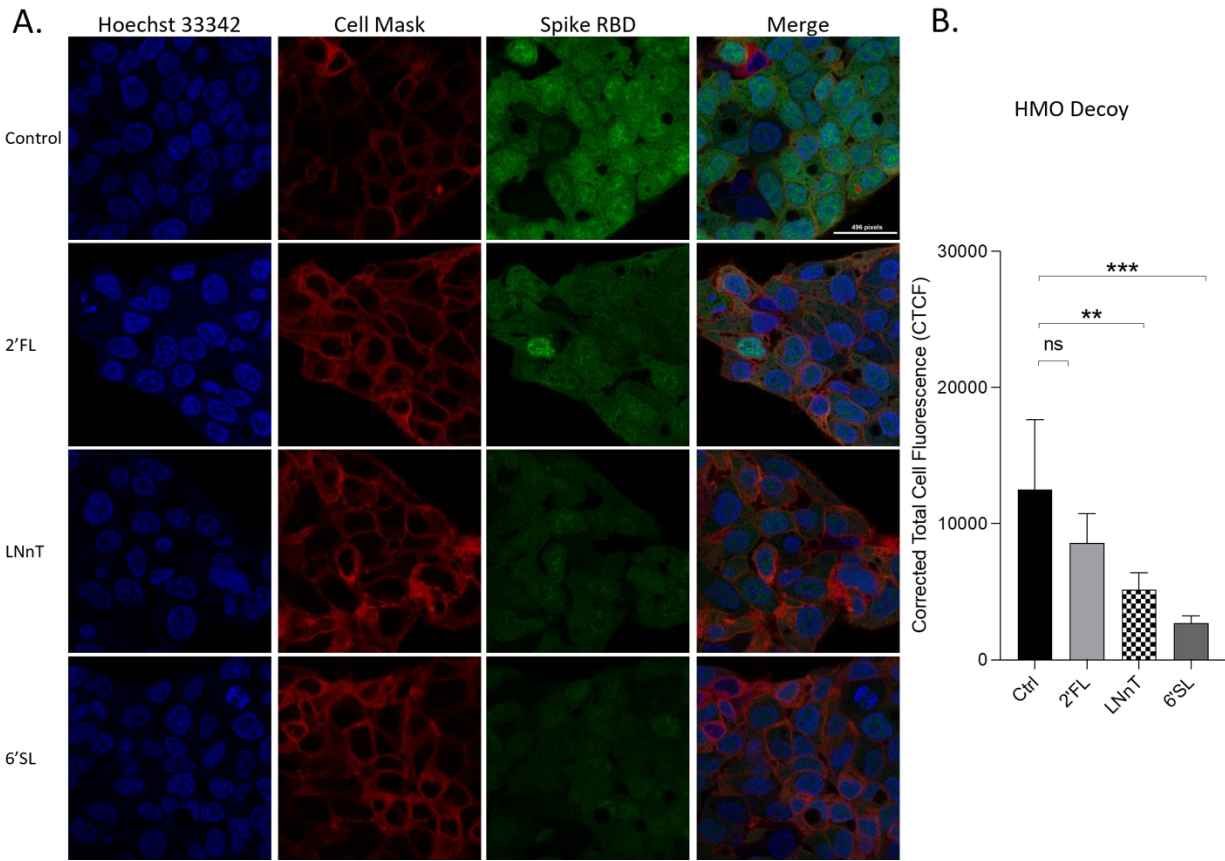
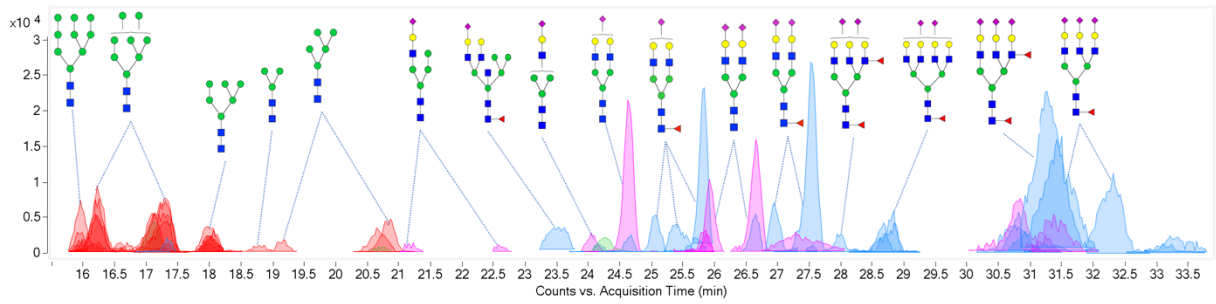


Figure 2.2. Inhibition of HMOs on the binding between HepG2 cells and Spike protein RBD.

Fluorescent labelled proteins were preincubated with 1 mg per mL 2'-FL, 6'-SL, and LNnT, respectively. The preincubation was performed at room temperature for 30 min. **(A)** Immunofluorescence for S protein RBD binding. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), S protein RBD, and merged image. Scale bar, 496 pixels. **(B)** Quantification of fluorescent intensity of Spike protein RBD binding. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the statistical significance between groups compared (**p < 0.01%; ***p < 0.001%; ns p < 0.05).

A. Chromatogram of N-Glycome of HepG2 cells



B. N-Glycoforms from ACE2 Protein

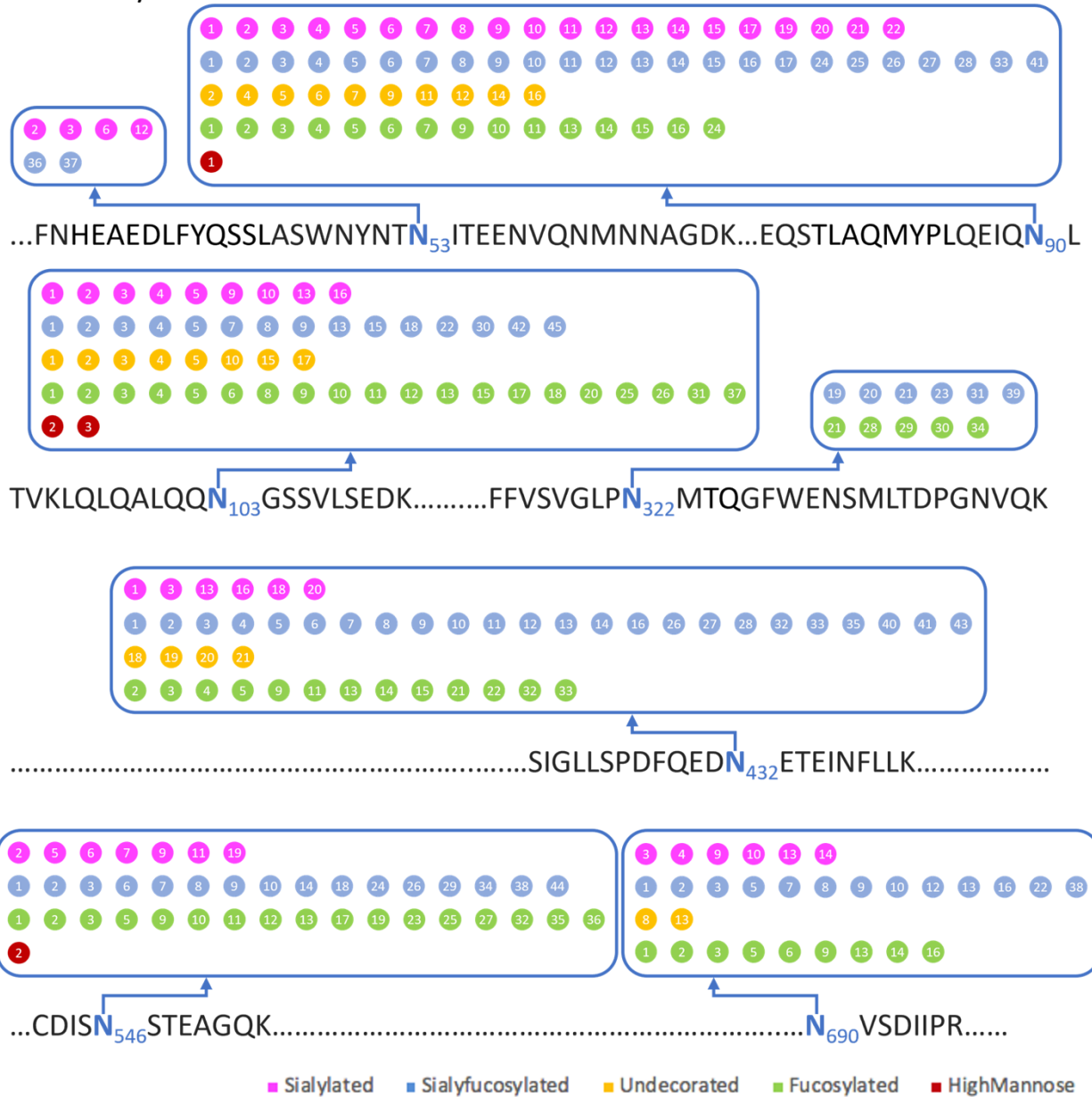
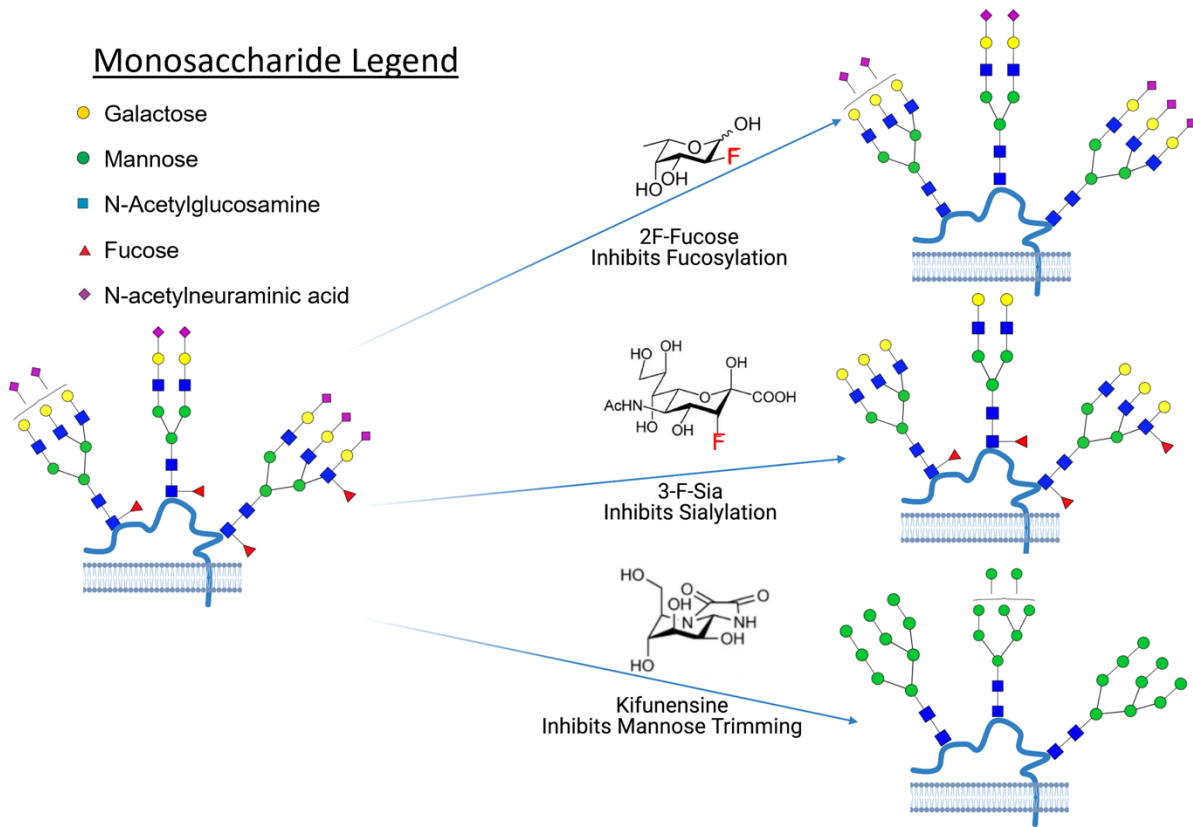


Figure 2.3. Cell membrane N-glycome and site-specific occupancy of ACE2 receptor in HepG2 cells. (A) Individual N-glycan species of HepG2 host cells. LC-MS peaks were color coded to assign glycan subtype. Abundant peaks are annotated with putative structures. Symbol nomenclature is used for representing glycan structures (<https://www.ncbi.nlm.nih.gov/glycans/snfg.html>). (B) Site-specific occupancy of ACE2 receptor in HepG2 cells. The N-glycoforms from ACE2 protein extracted from HepG2 cells are distributed on 7 glycosites. The labeled numbers inside dots denote identified individual glycan and the details were shown in Table S2.

A. Schematic Representation of Modifying Host Cell Surface Glycosylation



B. Relative Abundance of N-Glycan Subtypes from Modified HepG2 Cells

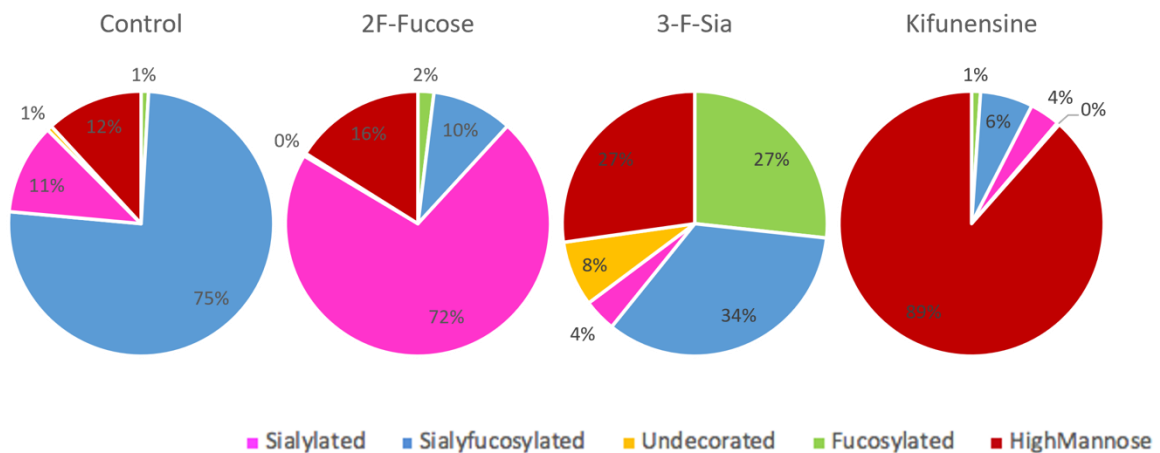


Figure 2.4. Host Cell Surface Glycome Modification. 2F-Fucose (Fucosyltransferase Inhibitor); 3-F-Sia (Sialyltransferase Inhibitor). **(A)** Metabolic engineering strategy for altering host cell glycosylation. Symbol nomenclature is used for representing glycan structures (<https://www.ncbi.nlm.nih.gov/glycans/snfg.html>). **(B)** N-Glycome Profiles of unmodified and modified HepG2 cells from LC-MS analysis. Compound list and details are shown in Table S3. Pie charts were color coded to assign glycan subtype. Numbers inside pie charts denote the relative abundance of each identified glycan subtype.

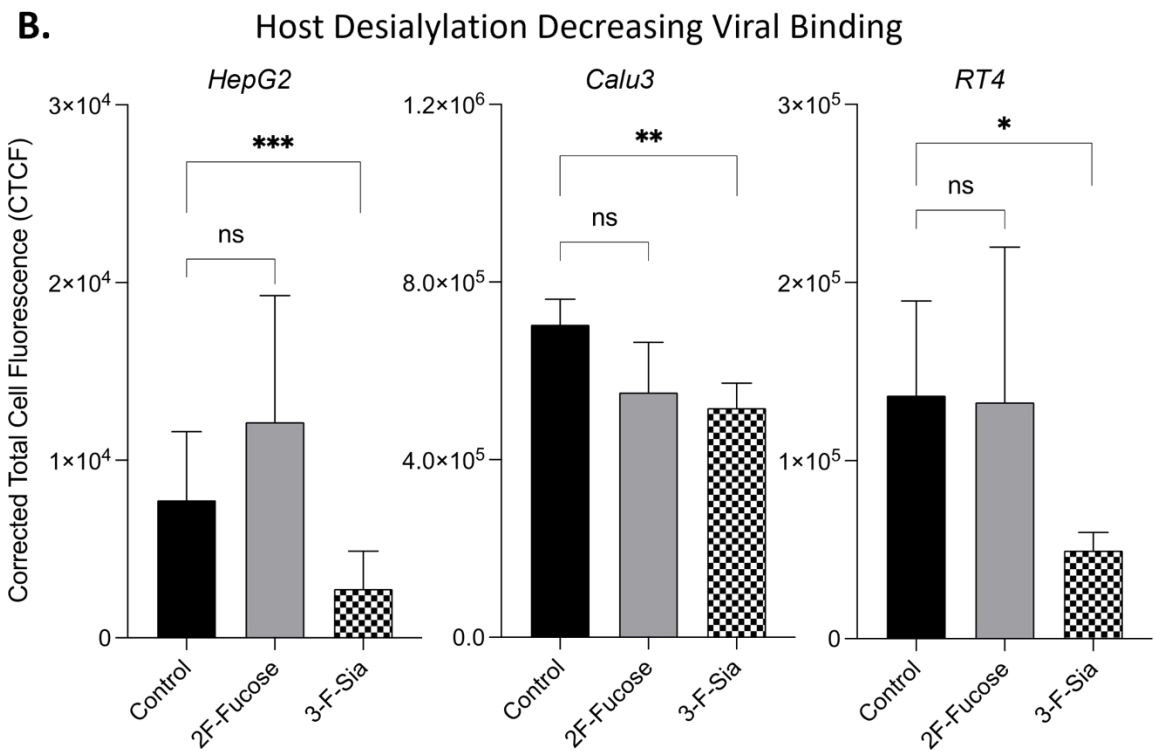
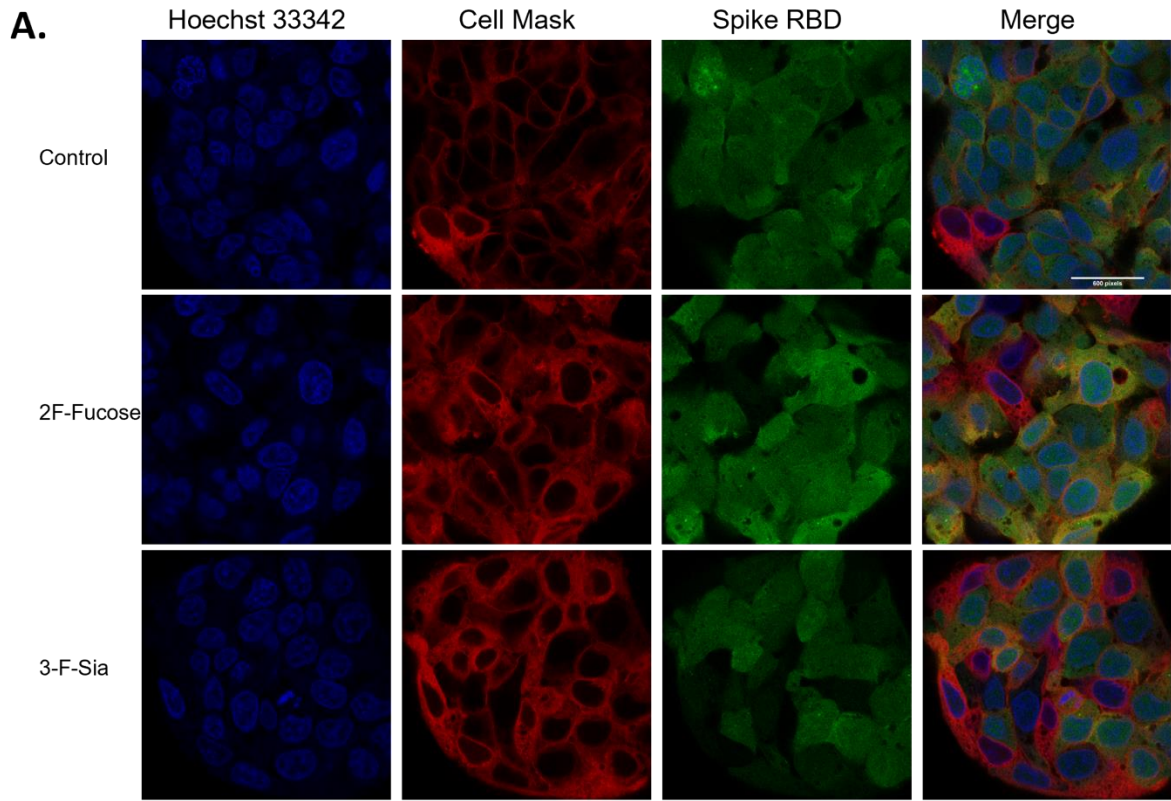


Figure 2.5. Remodeling host glycome alters binding between host cells and Spike protein RBD.

2F-Fucose (Fucosyltransferase Inhibitor); 3-F-Sia (Sialyltransferase Inhibitor). **(A)** Immunofluorescence for S protein RBD binding with modified cells. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), S protein RBD, and merged image. Scale bar, 600 pixels. **(B)** Quantification of fluorescent intensity of Spike protein RBD binding. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the statistical significance between groups compared (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns $p < 0.05$).

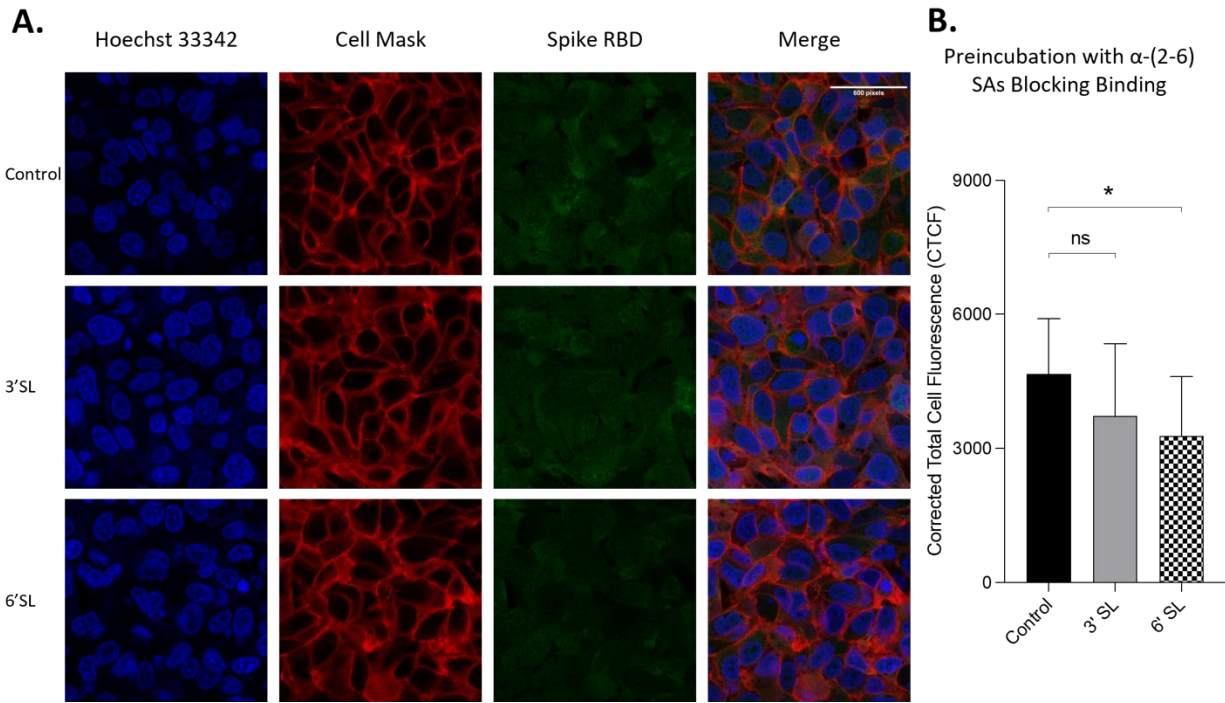


Figure 2.6. Sialylated 6'-SL HMOs inhibited binding of Spike protein RBD to HepG2 Cells.

Fluorescent labelled proteins were preincubated with 1 mg per mL 3'-FL(3'-sialyllactose) and 6'-SL(6'-sialyllactose) respectively at room temperature. **(A)** Immunofluorescence for S protein RBD binding with cells. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), S protein RBD, and merged image. Scale bar, 600 pixels. **(B)** Quantification of fluorescent intensity of Spike protein RBD binding. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the statistical significance between groups compared (* $p < 0.05$; ns $p < 0.05$).

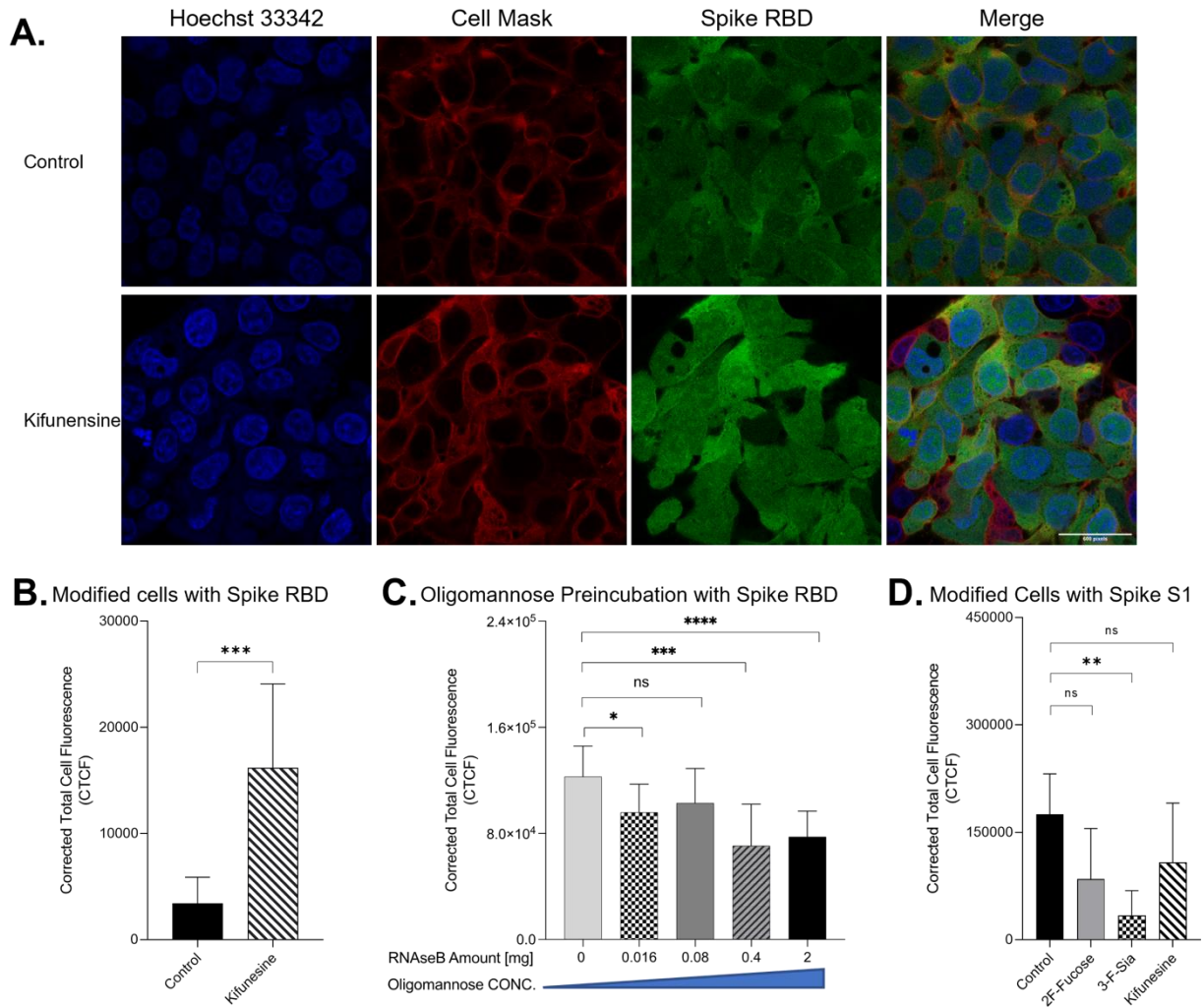


Figure 2.7. Introducing high mannose glycans into viral binding. (A) Immunofluorescence for S protein RBD binding with modified cells. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), S protein RBD, and merged image. Scale bar, 600 pixels. Quantification of fluorescent intensity of Spike protein RBD (B) (C) or S1 subunit (D) binding. (C) Fluorescent labelled proteins were preincubated with purified high mannose at room temperature for 30 min before binding. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the

statistical significance between groups compared (*p< 0.05%; **p< 0.01%; ***p< 0.001%;
****p< 0.0001%; ns p<0.05).

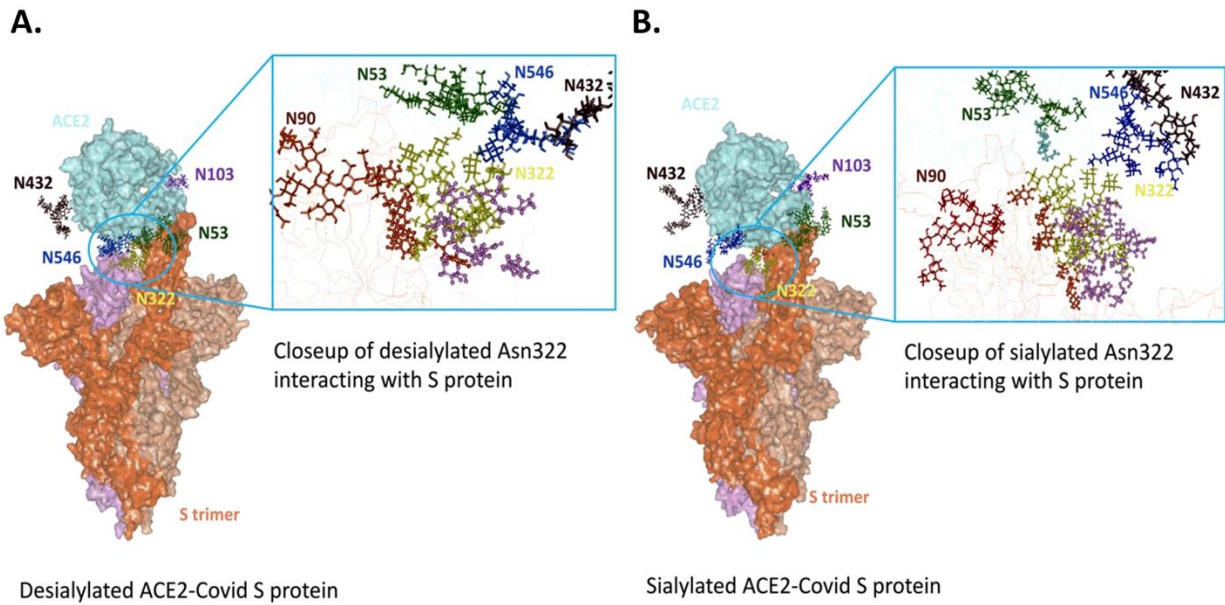


Figure 2.8. Modelled sialylated and desialylated ACE2-Covid S protein complexes. 3D structural modeling of glycosylated ACE2 interacting with S-protein. Results from glycomics and glycoproteomics of HEPG2 cell lines were used to generate **(A)** fully-desialylated and **(B)** fully-sialylated homologs of ACE2, interacting with S-protein.

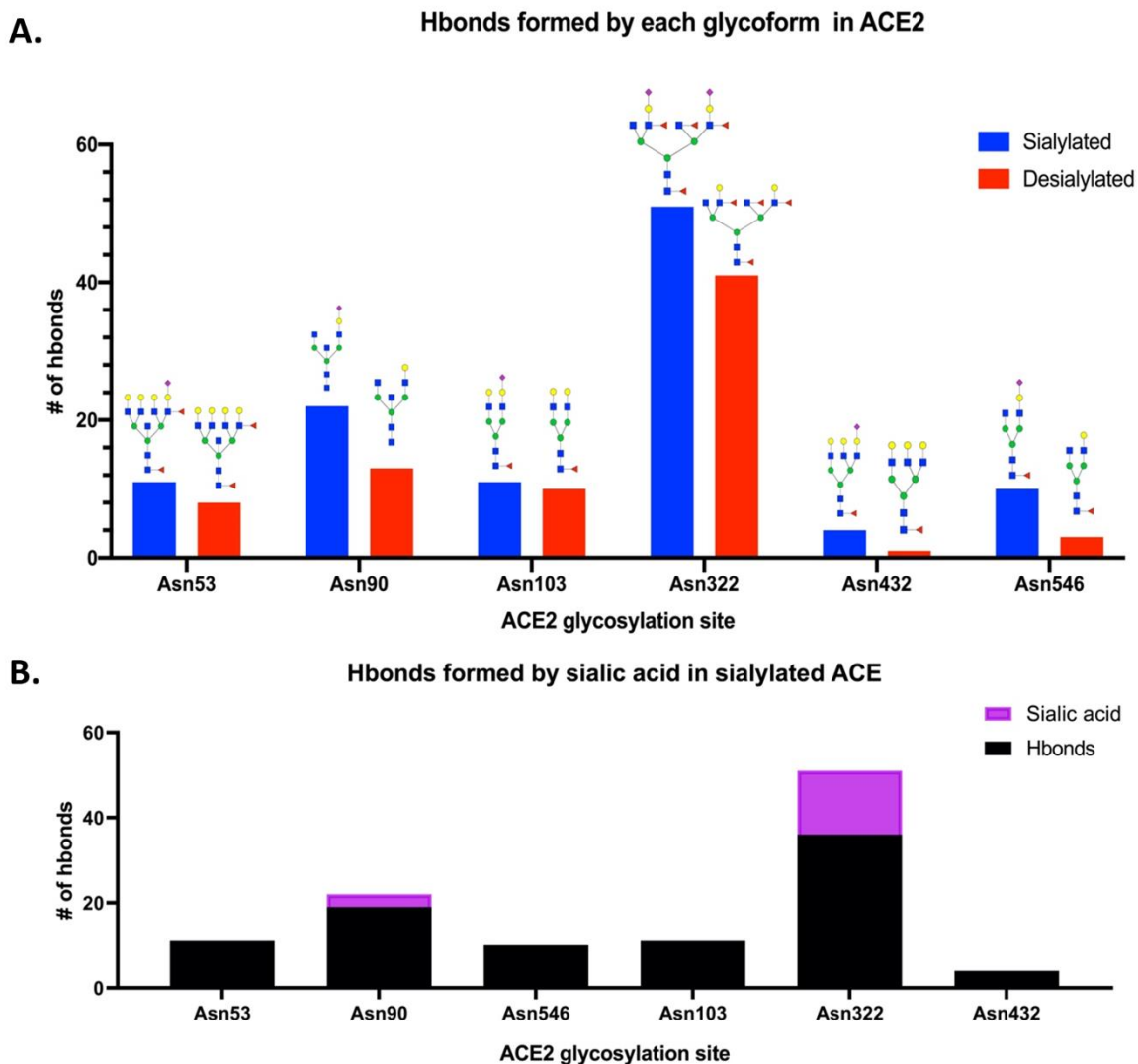


Figure 2.9. Interactions of glycosylated ACE2 and S-protein were revealed using molecular dynamics simulations. (A) The number of intramolecular hydrogen bonds was drastically higher for each fully-sialylated N-glycan compared to the desialylated glycoform. **(B)** For Asn 90 and Asn 322, the sialic acid residue in the glycoform accounted for ~10% of hydrogen bonds.

TABLES

Table 2.1. Summary of Glycoproteomic Profiles of ACE2 proteins. The number of glycoforms was shown in Table 1.

Subtype of N-Glycans detected on glycosites	Fucosylated	Sialyfucosylated	Undecorated	Sialylated	HM
Recombinant Human ACE2	147	175	68	119	16
ACE2 from HepG2 cells	100	120	38	65	5

REFERENCE

1. N. Zhu *et al.*, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-733 (2020).
2. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **183**, 1735 (2020).
3. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J Virol* **94**, (2020).
4. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
5. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e278 (2020).
6. C. Gstöttner *et al.*, Structural and Functional Characterization of SARS-CoV-2 RBD Domains Produced in Mammalian Cells. *Anal Chem* **93**, 6839-6847 (2021).
7. J. S. Peiris *et al.*, Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319-1325 (2003).
8. C. S. G. o. t. I. C. o. T. o. Viruses, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544 (2020).
9. W. Li *et al.*, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).
10. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **5**, 562-569 (2020).
11. W. Tai *et al.*, Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol Immunol* **17**, 613-620 (2020).
12. J. Lan *et al.*, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220 (2020).
13. X. Tian *et al.*, Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg Microbes Infect* **9**, 382-385 (2020).
14. I. Hamming *et al.*, Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* **203**, 631-637 (2004).
15. D. Park *et al.*, Enterocyte glycosylation is responsive to changes in extracellular conditions: implications for membrane functions. *Glycobiology* **27**, 847-860 (2017).
16. D. Park *et al.*, Characteristic Changes in Cell Surface Glycosylation Accompany Intestinal Epithelial Cell (IEC) Differentiation: High Mannose Structures Dominate the Cell Surface Glycome of Undifferentiated Enterocytes. *Mol Cell Proteomics* **14**, 2910-2921 (2015).
17. A. Shajahan *et al.*, Comprehensive characterization of N- and O- glycosylation of SARS-CoV-2 human receptor angiotensin converting enzyme 2. *Glycobiology* **31**, 410-424 (2021).

18. Y. Watanabe, J. D. Allen, D. Wrapp, J. S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330-333 (2020).
19. A. Shajahan, N. T. Supekar, A. S. Gleinich, P. Azadi, Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* **30**, 981-988 (2020).
20. A. R. Mehdipour, G. Hummer, Dual nature of human ACE2 glycosylation in binding to SARS-CoV-2 spike. *Proc Natl Acad Sci U S A* **118**, (2021).
21. P. Zhao *et al.*, Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. *Cell Host Microbe* **28**, 586-601.e586 (2020).
22. D. D. Eric W. Stawiski , Kushal Suryamohan , Ravi Gupta , Frederic A. Fellouse , J. Fah Sathirapongsasuti , Jiang Liu , Ying-Ping Jiang , Aakrosh Ratan , Monika Mis , Devi Santhosh , Sneha Somasekar , Sangeetha Mohan , Sameer Phalke , Boney Kuriakose , Aju Antony , Jagath R. Junutula , Stephan C. Schuster , Natalia Jura , Somasekar Seshagiri. (<https://www.biorxiv.org/>, 2021).
23. K. K. Chan *et al.*, Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **369**, 1261-1265 (2020).
24. S. Y. Kim *et al.*, Characterization of heparin and severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) spike glycoprotein binding interactions. *Antiviral Res* **181**, 104873 (2020).
25. T. M. Clausen *et al.*, SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. *Cell* **183**, 1043-1057.e1015 (2020).
26. R. E. Moore, L. L. Xu, S. D. Townsend, Prospecting Human Milk Oligosaccharides as a Defense Against Viral Infections. *ACS Infect Dis*, (2021).
27. S. Wu, R. Grimm, J. B. German, C. B. Lebrilla, Annotation and structural analysis of sialylated human milk oligosaccharides. *J Proteome Res* **10**, 856-868 (2011).
28. S. Wu, N. Tao, J. B. German, R. Grimm, C. B. Lebrilla, Development of an annotated library of neutral human milk oligosaccharides. *J Proteome Res* **9**, 4138-4151 (2010).
29. Q. Li, Y. Xie, G. Xu, C. B. Lebrilla, Identification of potential sialic acid binding proteins on cell membranes by proximity chemical labeling. *Chem Sci* **10**, 6199-6209 (2019).
30. Q. Li, Y. Xie, M. Wong, M. Barboza, C. B. Lebrilla, Comprehensive structural glycomic characterization of the glycocalyxes of cells and tissues. *Nat Protoc* **15**, 2668-2704 (2020).
31. C. Xu *et al.*, Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Sci Adv* **7**, (2021).
32. S. J. Park *et al.*, CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **29**, 320-331 (2019).
33. O. Guvench *et al.*, CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* **7**, 3162-3180 (2011).
34. J. Huang *et al.*, CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **14**, 71-73 (2017).
35. B. Acun *et al.*, Scalable Molecular Dynamics with NAMD on the Summit System. *IBM J Res Dev* **62**, 1-9 (2018).

36. S. S. Mallajosyula, S. Jo, W. Im, A. D. MacKerell, Molecular dynamics simulations of glycoproteins using CHARMM. *Methods Mol Biol* **1273**, 407-429 (2015).
37. S. Jo *et al.*, CHARMM-GUI 10 years for biomolecular modeling and simulation. *J Comput Chem* **38**, 1114-1124 (2017).
38. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).
39. E. Castanys-Muñoz, M. J. Martin, P. A. Prieto, 2'-fucosyllactose: an abundant, genetically determined soluble glycan present in human milk. *Nutr Rev* **71**, 773-789 (2013).
40. A. Digre, C. Lindskog, The Human Protein Atlas-Spatial localization of the human proteome in health and disease. *Protein Sci* **30**, 218-233 (2021).
41. P. J. Thul, C. Lindskog, The human protein atlas: A spatial map of the human proteome. *Protein Sci* **27**, 233-244 (2018).
42. Q. Zhou, Y. Xie, M. Lam, C. B. Lebrilla, -Glycomic Analysis of the Cell Shows Specific Effects of Glycosyl Transferase Inhibitors. *Cells* **10**, (2021).
43. J. A. Villalobos, B. R. Yi, I. S. Wallace, 2-Fluoro-L-Fucose Is a Metabolically Incorporated Inhibitor of Plant Cell Wall Polysaccharide Fucosylation. *PLoS One* **10**, e0139091 (2015).
44. Y. Zhou *et al.*, Inhibition of fucosylation by 2-fluorofucose suppresses human liver cancer HepG2 cell proliferation and migration as well as tumor formation. *Sci Rep* **7**, 11563 (2017).
45. K. Suzuki, S. Daikoku, S. H. Son, Y. Ito, O. Kanie, Synthetic study of 3-fluorinated sialic acid derivatives. *Carbohydr Res* **406**, 1-9 (2015).
46. C. R. Varki A, Esko JD, *et al.*, editors., *Essentials of Glycobiology*, 3rd edition.
47. K. Shinya *et al.*, Avian flu: influenza virus receptors in the human airway. *Nature* **440**, 435-436 (2006).
48. J. M. Nicholls, A. J. Bourne, H. Chen, Y. Guan, J. S. Peiris, Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respir Res* **8**, 73 (2007).
49. M. Schneider, E. Al-Shareffi, R. S. Haltiwanger, Biological functions of fucose in mammals. *Glycobiology* **27**, 601-618 (2017).
50. B. Ma, J. L. Simala-Grant, D. E. Taylor, Fucosylation in prokaryotes and eukaryotes. *Glycobiology* **16**, 158R-184R (2006).
51. R. J. G. Hulswit *et al.*, Human coronaviruses OC43 and HKU1 bind to 9-. *Proc Natl Acad Sci U S A* **116**, 2681-2690 (2019).
52. D. P. Han, M. Lohani, M. W. Cho, Specific asparagine-linked glycosylation sites are critical for DC-SIGN- and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry. *J Virol* **81**, 12029-12039 (2007).
53. V. Kommineni *et al.*, In Vivo Glycan Engineering via the Mannosidase I Inhibitor (Kifunensine) Improves Efficacy of Rituximab Manufactured in. *Int J Mol Sci* **20**, (2019).
54. H. Y. Choi *et al.*, N-glycan Remodeling Using Mannosidase Inhibitors to Increase High-mannose Glycans on Acid α -Glucosidase in Transgenic Rice Cell Cultures. *Sci Rep* **8**, 16130 (2018).
55. L. Nguyen *et al.*, Sialic acid-Dependent Binding and Viral Entry of SARS-CoV-2. *bioRxiv*, 2021.2003.2008.434228 (2021).

56. S. K. Wong, W. Li, M. J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J Biol Chem* **279**, 3197-3201 (2004).
57. R. Trebbien, L. E. Larsen, B. M. Viuff, Distribution of sialic acid receptors and influenza A virus of avian and swine origin in experimentally infected pigs. *Virology* **8**, 434 (2011).
58. N. H. Wu, F. Meng, M. Seitz, P. Valentin-Weigand, G. Herrler, Sialic acid-dependent interactions between influenza viruses and *Streptococcus suis* affect the infection of porcine tracheal cells. *J Gen Virol* **96**, 2557-2568 (2015).
59. C. Sieben, E. Sezgin, C. Eggeling, S. Manley, Influenza A viruses use multivalent sialic acid clusters for cell binding and receptor activation. *PLoS Pathog* **16**, e1008656 (2020).
60. R. J. Almond *et al.*, Differential immunogenicity and allergenicity of native and recombinant human lactoferrins: role of glycosylation. *Eur J Immunol* **43**, 170-181 (2013).
61. D. Morniroli *et al.*, The Antiviral Properties of Human Milk: A Multitude of Defence Tools from Mother Nature. *Nutrients* **13**, (2021).
62. J. T. Smilowitz *et al.*, Human milk secretory immunoglobulin a and lactoferrin N-glycans are altered in women with gestational diabetes mellitus. *J Nutr* **143**, 1906-1912 (2013).
63. J. M. Jorgensen *et al.*, Lipid-Based Nutrient Supplements During Pregnancy and Lactation Did Not Affect Human Milk Oligosaccharides and Bioactive Proteins in a Randomized Trial. *J Nutr* **147**, 1867-1874 (2017).
64. X. Wang, Y. Zhou, N. Jiang, Q. Zhou, W. L. Ma, Persistence of intestinal SARS-CoV-2 infection in patients with COVID-19 leads to re-admission after pneumonia resolved. *Int J Infect Dis* **95**, 433-435 (2020).
65. G. C. Hansson, Mucus and mucins in diseases of the intestinal and respiratory tracts. *J Intern Med* **285**, 479-490 (2019).
66. D. T. Tran, K. G. Ten Hagen, Mucin-type O-glycosylation during development. *J Biol Chem* **288**, 6921-6929 (2013).
67. L. E. Tailford, E. H. Crost, D. Kavanaugh, N. Juge, Mucin glycan foraging in the human gut microbiome. *Front Genet* **6**, 81 (2015).
68. C. M. Radcliffe *et al.*, Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. *J Biol Chem* **282**, 7405-7415 (2007).
69. U. Möglinger *et al.*, Alterations of the Human Skin. *Front Oncol* **8**, 70 (2018).
70. H. J. An *et al.*, Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol Cell Proteomics* **11**, M111.010660 (2012).
71. C. I. Balog *et al.*, N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol Cell Proteomics* **11**, 571-585 (2012).
72. L. R. Ruhaak *et al.*, Differential N-Glycosylation Patterns in Lung Adenocarcinoma Tissue. *J Proteome Res* **14**, 4538-4549 (2015).
73. D. Park *et al.*, *Salmonella Typhimurium* Enzymatically Landscapes the Host Intestinal Epithelial Cell (IEC) Surface Glycome to Increase Invasion. *Mol Cell Proteomics* **15**, 3653-3664 (2016).
74. P. C. Reading, J. L. Miller, E. M. Anders, Involvement of the mannose receptor in infection of macrophages by influenza virus. *J Virol* **74**, 5190-5197 (2000).

75. J. P. Upham, D. Pickett, T. Irimura, E. M. Anders, P. C. Reading, Macrophage receptors for influenza A virus: role of the macrophage galactose-type lectin and mannose receptor in viral entry. *J Virol* **84**, 3730-3737 (2010).
76. J. L. Miller *et al.*, The mannose receptor mediates dengue virus infection of macrophages. *PLoS Pathog* **4**, e17 (2008).
77. D. G. Nguyen, J. E. Hildreth, Involvement of macrophage mannose receptor in the binding and transmission of HIV by macrophages. *Eur J Immunol* **33**, 483-493 (2003).
78. N. K. Routhu *et al.*, Glycosylation of Zika Virus is Important in Host-Virus Interaction and Pathogenic Potential. *Int J Mol Sci* **20**, (2019).
79. A. Varki *et al.* (2015).
80. G. Roberts, E. Tarelli, K. A. Homer, J. Philpott-Howard, D. Beighton, Production of an endo-beta-N-acetylglucosaminidase activity mediates growth of *Enterococcus faecalis* on a high-mannose-type glycoprotein. *J Bacteriol* **182**, 882-890 (2000).
81. M. Wong *et al.*, Metabolic flux analysis of the neural cell glycoalyx reveals differential utilization of monosaccharides. *Glycobiology* **30**, 859-871 (2020).

SUPPLEMENTARY INFORMATION

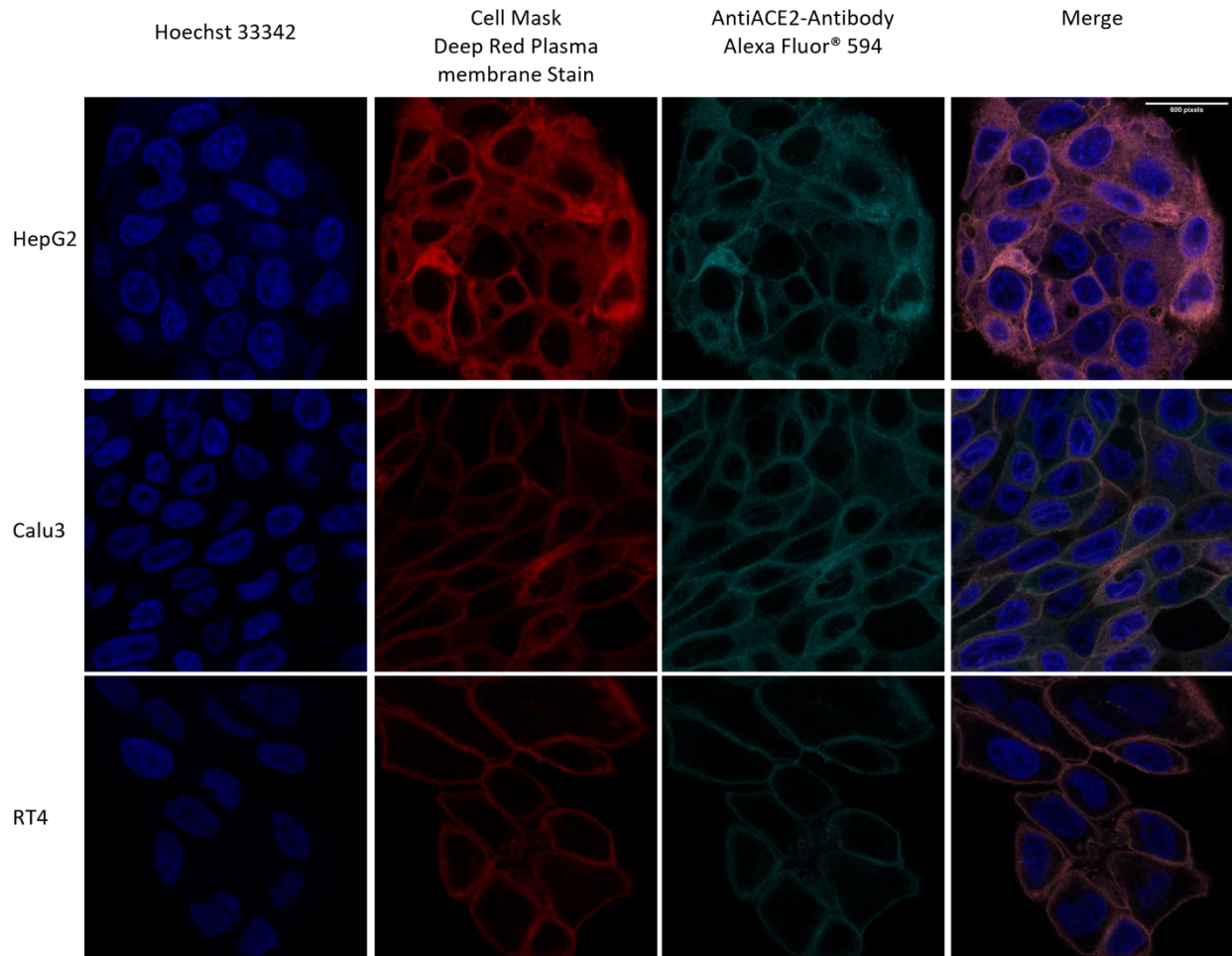


Figure. S2.1 Validating the expression of ACE2 in HepG2 cells using immunofluorescence. Cells were incubated with mouse monoclonal antibody. The images show HepG2, Calu3 and RT4 cells expressing huACE2. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), anti-ACE2 antibody, and merged image. Scale bar, 600 pixels.

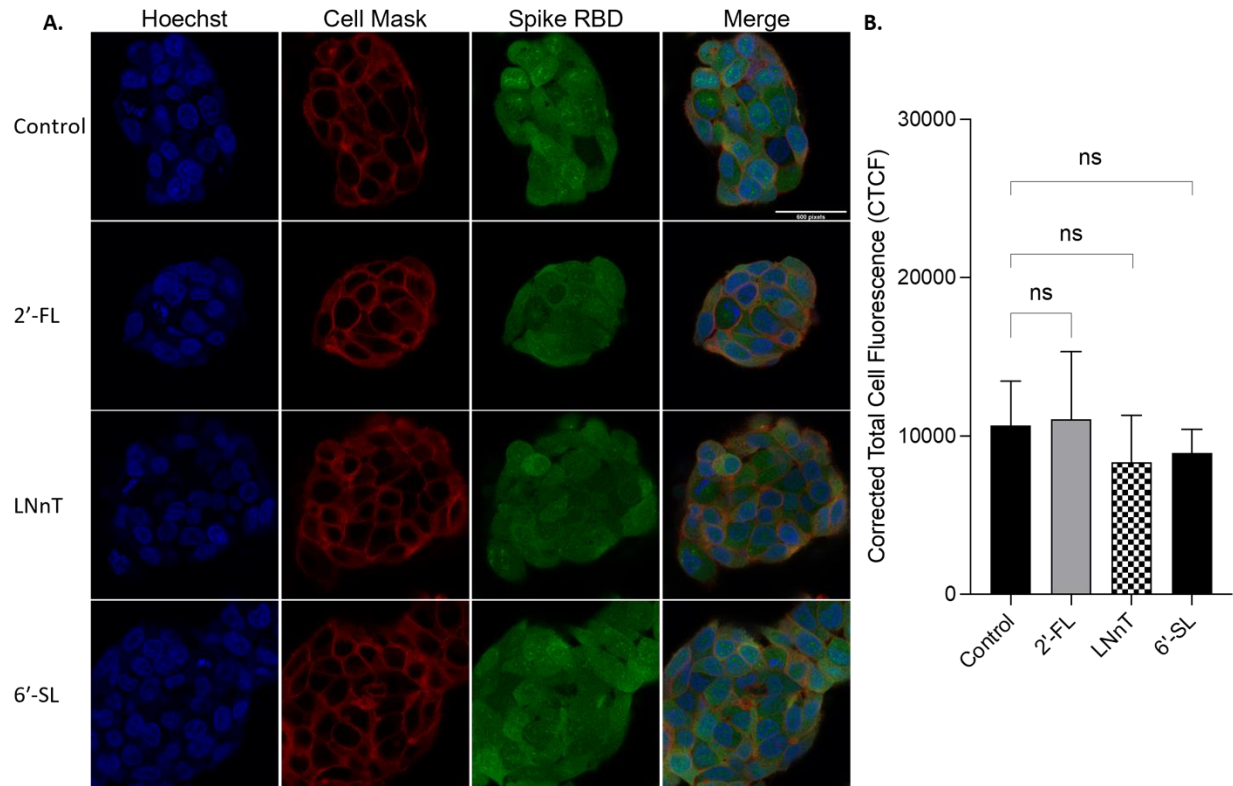


Figure S2.2. Effect of HMOs on the binding between HepG2 cells and Spike protein RBD. HepG2 cells were preincubated with 1 mg per mL 2'-FL, 6'-SL, and LNnT, respectively. After 30 min-preincubation at room temperature, cells were washed with PBS to remove exogenous decoys before fixation. **(A)** Immunofluorescence for S protein RBD binding with HepG2 cells. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), S protein RBD, and merged image. Scale bar, 600 pixels. **(B)** Quantification of fluorescent intensity of Spike protein RBD binding. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the statistical significance between groups compared (ns $p < 0.05$).

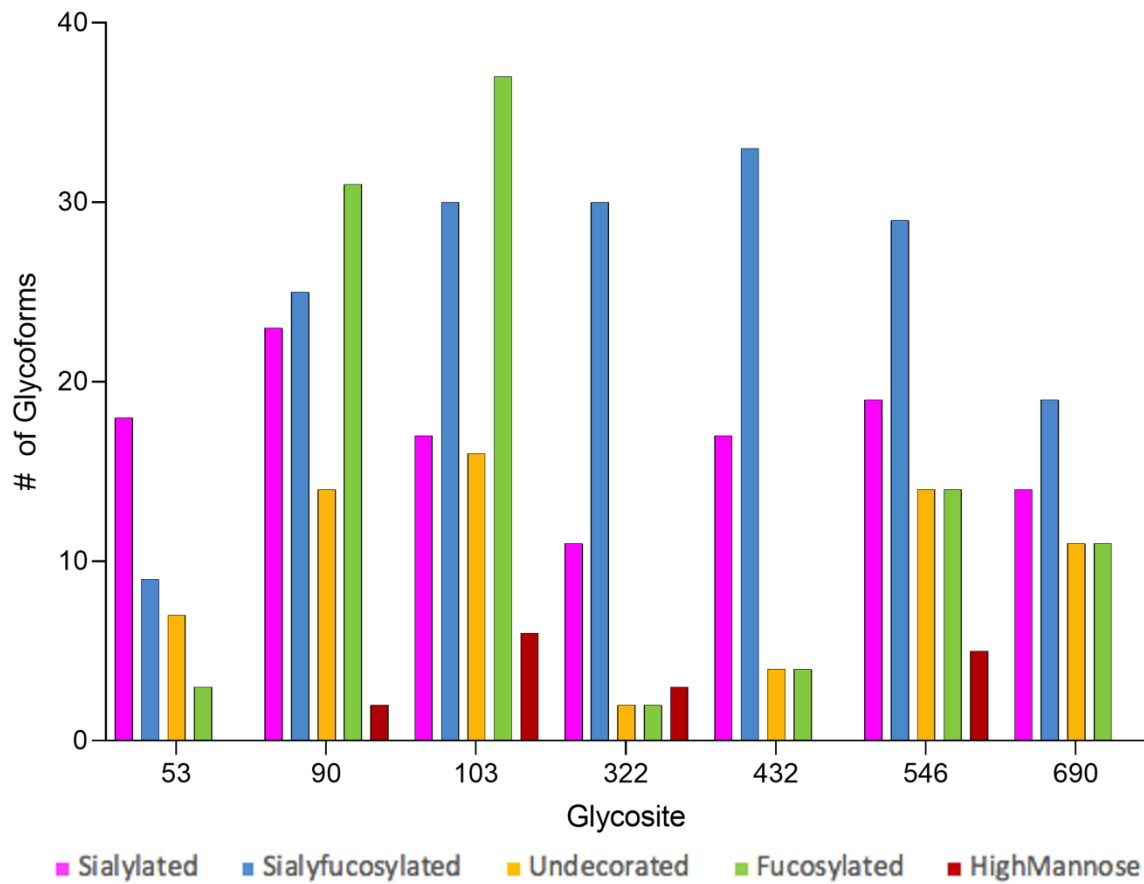


Figure S2.3. Site-specific occupancy of recombinant ACE2 receptor. The number of glycoforms for each glycosylation site of recombinant protein ACE2 derived from HEK293 cells.

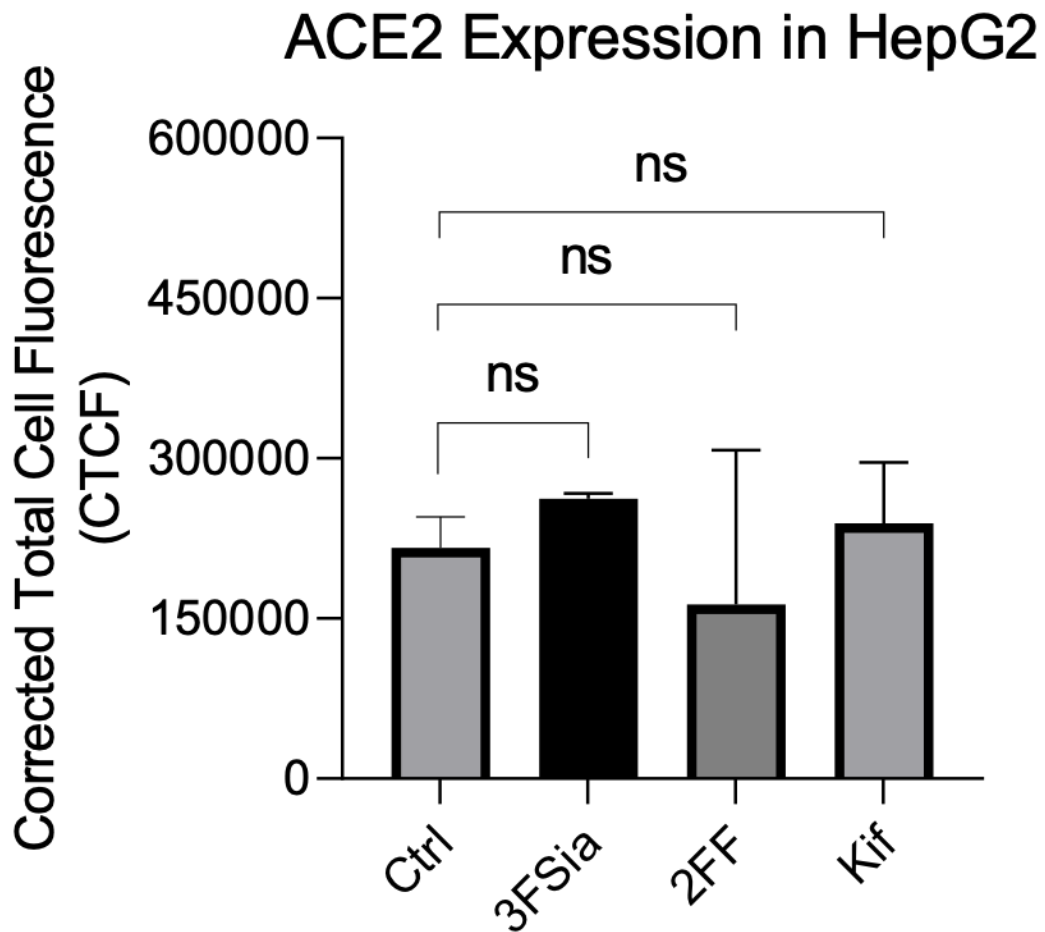


Figure S2.4. Host cell glycolyx remodeling not altering ACE2 expression on HepG2 cells.

Quantification of fluorescent intensity of anti-ACE2 antibodies on HepG2 cells. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the statistical significance between groups compared (ns $p < 0.05$).

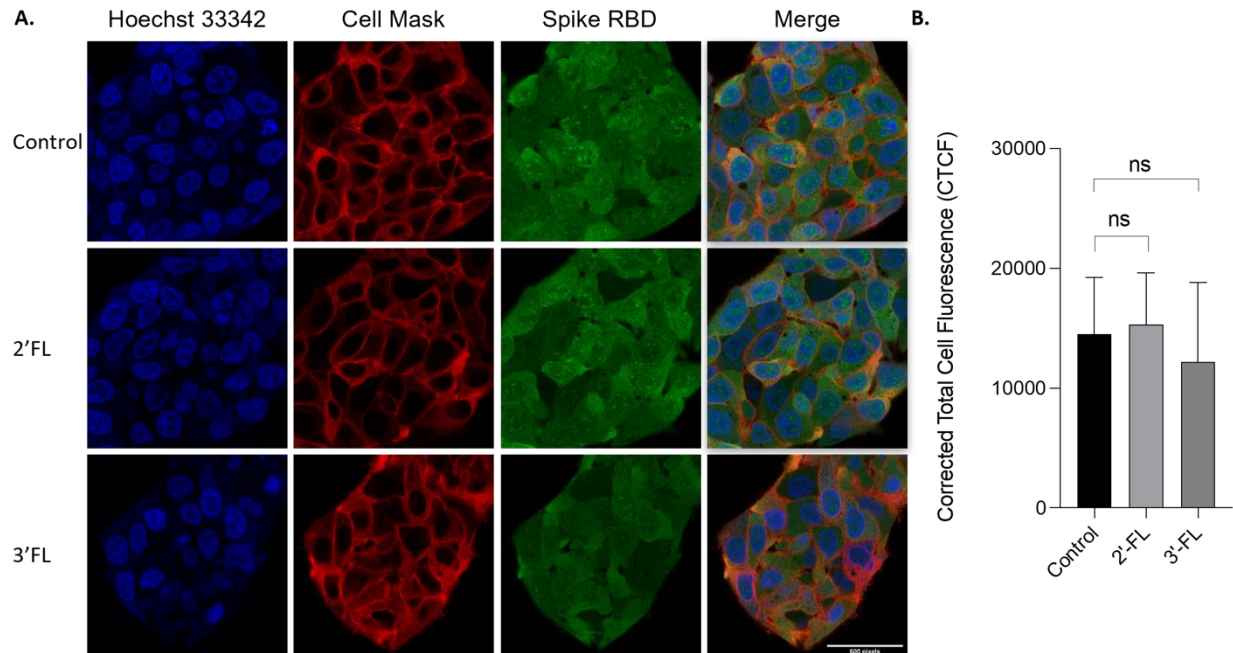


Figure S2.5. Effect of fucosylated HMOs on the binding between HepG2 cells and Spike protein RBD. S protein RBD with 2'-FL or 3'-FL for 30 min at room temperature. **(A)** Immunofluorescence for S protein RBD binding with HepG2 cells. The columns (from left to right) show staining of nuclear acid (Hoechst 33342), plasma membrane (CellMask™ Deep Red), S protein RBD, and merged image. Scale bar, 600 pixels. **(B)** Quantification of fluorescent intensity of Spike protein RBD binding. Fluorescence intensity was quantified for selected cell area. Quantification was performed with software ImageJ. Asterisks indicate the statistical significance between groups compared (ns $p < 0.05$).

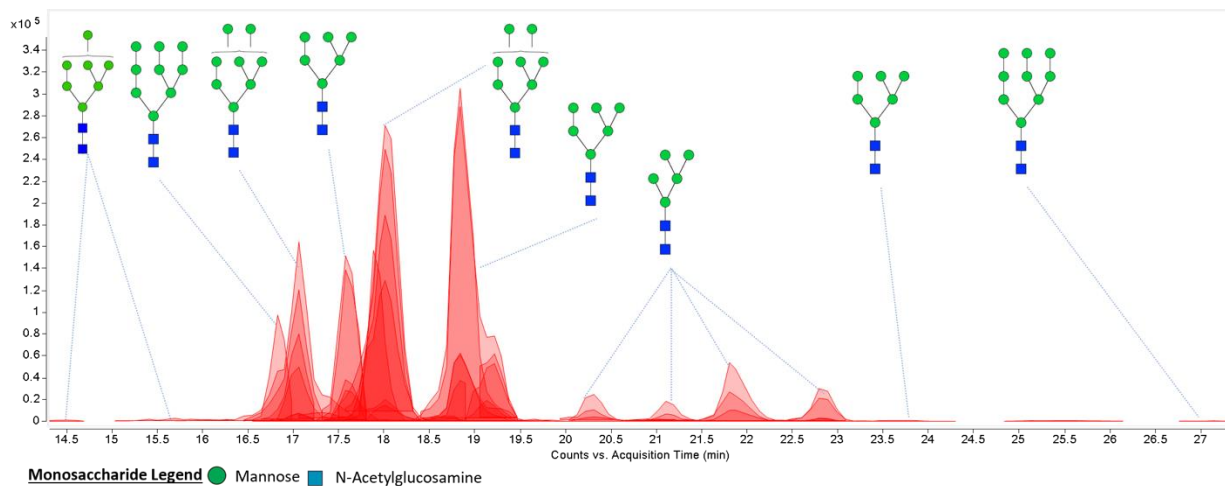


Figure S2.6. Chromatogram of high mannose N-glycans released from RNase B. Abundant peaks are annotated with putative structures. Symbol nomenclature is used for representing glycan structures (<https://www.ncbi.nlm.nih.gov/glycans/snfg.html>).

Table S2.1. Glycoproteomic information of ACE2 derived from HepG2 cells

Peptide Sequence	Modification	Modified Amino Acid	Modified Position	Glycans	Label in Figure 2. 3.	Relative Abundance	Glycan Subtype
FNHEAEDLFYQSSLASWNYN TNITEENVQNMNNAGDK	NGlycan/3139.1366	N	53	HexNAc(7)Hex(7)Fuc(2)NeuAc(1)	36	0.0326	Sialyfucoylated
FNHEAEDLFYQSSLASWNYN TNITEENVQNMNNAGDK	NGlycan/2992.0834	N	53	HexNAc(5)Hex(5)Fuc(4)NeuAc(2)	37	0.0298	Sialyfucoylated
FNHEAEDLFYQSSLASWNYN TNITEENVQNMNNAGDK	NGlycan/1954.7036	N	53	HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	2	0.0235	Sialylated
FNHEAEDLFYQSSLASWNYN TNITEENVQNMNNAGDK	NGlycan/2157.7829	N	53	HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	12	0.0227	Sialylated
FNHEAEDLFYQSSLASWNYN TNITEENVQNMNNAGDK	NGlycan/2319.8358 ; Deamidated/0.984 0	N,N	53,58	HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	6	0.00539	Sialylated
FNHEAEDLFYQSSLASWNYN TNITEENVQNMNNAGDK	NGlycan/2116.7564	N	53	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	0.0253	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2059.7349	N	90	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	0.592	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2059.7349	E,N	75,90	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	0.0441	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2350.8304	N	90	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	0.275	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2350.8304	E,N	75,90	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	0.0259	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1897.6821	N	90	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.7	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1897.6821	E,N	75,90	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.0349	Sialyfucoylated

EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/1897.6821	Q,N	76,90	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.0126	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2367.8457	N	90	HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	4	0.0617	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2367.8457	E,N	75,90	HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	4	0.0223	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2205.7928	N	90	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	5	0.0237	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2789.9993	N	90	HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	6	0.328	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2424.8671	N	90	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	7	0.218	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2424.8671	E,N	75,90	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	7	0.0155	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2715.9625	N	90	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	8	0.221	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2715.9625	E,N	75,90	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	8	0.0392	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2262.8143	N	90	HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	9	0.212	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2100.7615	N	90	HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	10	0.149	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/3081.0947	N	90	HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	11	0.0402	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2732.9779	E,N	75,90	HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	12	0.177	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2732.9779	N	90	HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	12	0.127	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/3007.0580	N	90	HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	13	0.0612	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2553.9097	N	90	HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	14	0.153	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2553.9097	E,N	75,90	HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	14	0.0247	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2221.7878	N	90	HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	15	0.0205	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/3024.0733	N	90	HexNAc(5)Hex(7)Fuc(2)NeuAc(2)	16	0.034	Sialyfucoylated

EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/3082.1151	E,N	75,90	HexNAc(6)Hex(7)Fuc(3)NeuAc(1)	17	0.124	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2465.8937	N	90	HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	24	0.0757	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2774.0044	E,N	75,90	HexNAc(6)Hex(6)Fuc(2)NeuAc(1)	25	0.075	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2627.9465	N	90	HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	26	0.0389	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/3372.1902	N	90	HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	27	0.0246	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2919.0419	N	90	HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	28	0.0411	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2570.9250	N	90	HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	33	0.0272	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2570.9250	E,N	75,90	HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	33	0.0225	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/3210.1373	N	90	HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	41	0.0106	Sialyfucoylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2075.7298	N	90	HexNAc(4)Hex(6)Fuc(0)NeuAc(1)	1	0.0551	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1954.7036	N	90	HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	2	2.02	Sialylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1954.7036	E,N	75,90	HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	2	0.123	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2116.7564	N	90	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	2	Sialylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2116.7564	E,N	75,90	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	0.11	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1913.6770	N	90	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.51	Sialylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1913.6770	E,N	75,90	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.0556	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2407.8518	N	90	HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	5	0.748	Sialylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2407.8518	E,N	75,90	HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	5	0.134	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2319.8358	N	90	HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	6	0.587	Sialylated

EQSTLAQMYPLQEIQNLTVK	NGlycan/2772.9840	N	90	HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	7	0.552	Sialylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2772.9840	E,N	75,90	HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	7	0.0723	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2481.8886	N	90	HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	8	0.49	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1751.6242	N	90	HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	9	0.47	Sialylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1751.6242	E,N	75,90	HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	9	0.0279	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2204.7724	N	90	HexNAc(4)Hex(5)Fuc(0)NeuAc(2)	10	0.125	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2610.9312	N	90	HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	11	0.437	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2157.7829	N	90	HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	12	0.418	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2278.8092	N	90	HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	13	0.136	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2569.9046	N	90	HexNAc(5)Hex(6)Fuc(0)NeuAc(2)	14	0.122	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2976.0634	N	90	HexNAc(7)Hex(6)Fuc(0)NeuAc(2)	15	0.0988	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2731.9575	N	90	HexNAc(5)Hex(7)Fuc(0)NeuAc(2)	17	0.0226	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1548.5448	N	90	HexNAc(3)Hex(4)Fuc(0)NeuAc(1)	19	0.0133	Sialylated
EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/2440.8620	Q,N	89,90	HexNAc(5)Hex(7)Fuc(0)NeuAc(1)	20	0.0193	Sialylated
EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/2440.8620	Q,N	81,90	HexNAc(5)Hex(7)Fuc(0)NeuAc(1)	20	0.0181	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2935.0368	N	90	HexNAc(6)Hex(7)Fuc(0)NeuAc(2)	21	0.0184	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2861.0001	N	90	HexNAc(5)Hex(6)Fuc(0)NeuAc(3)	22	0.0142	Sialylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1444.5339	N	90	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.112	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1606.5867	N	90	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	0.428	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1768.6395	N	90	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	0.251	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2076.7502	N	90	HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	4	0.297	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2133.7717	E,N	75,90	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	1.13	Fucosylated

EQSTLAQMYPLQEIQNLTVK	NGlycan/2133.7717	N	90	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	0.0772	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1914.6974	N	90	HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	6	0.315	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2644.9618	E,N	75,90	HexNAc(6)Hex(7)Fuc(2)NeuAc(0)	7	0.365	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1809.6661	N	90	HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	9	0.224	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1752.6446	N	90	HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	10	0.0827	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2279.8296	E,N	75,90	HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	11	0.334	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2279.8296	N	90	HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	11	0.101	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1971.7189	N	90	HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	13	0.136	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1971.7189	E,N	75,90	HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	13	0.0362	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2441.8824	N	90	HexNAc(5)Hex(7)Fuc(2)NeuAc(0)	14	0.0567	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/2117.7768	E,N	75,90	HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	15	0.138	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2117.7768	N	90	HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	15	0.0945	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/2117.7768	Q,N	86,90	HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	15	0.0191	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1930.6923	N	90	HexNAc(4)Hex(6)Fuc(1)NeuAc(0)	16	0.135	Fucosylated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1930.6923	E,N	75,90	HexNAc(4)Hex(6)Fuc(1)NeuAc(0)	16	0.0975	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2092.7452	N	90	HexNAc(4)Hex(7)Fuc(1)NeuAc(0)	24	0.0493	Fucosylated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1622.5816	N	90	HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	2	0.455	Undecorated
EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/1622.5816	Q,N	86,90	HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	2	0.181	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1663.6082	N	90	HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	4	1.22	Undecorated

EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/1663.6082	Q,N	89,90	HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	4	0.015	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1825.6610	N	90	HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	5	0.659	Undecorated
EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/1825.6610	Q,N	86,90	HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	5	0.0298	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1501.5553	N	90	HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	6	0.433	Undecorated
EQSTLAQMYPLQEIQNLTVK	Deamidated/0.984 0; NGlycan/1501.5553	Q,N	89,90	HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	6	0.00379	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1460.5288	N	90	HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	7	0.351	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1866.6875	N	90	HexNAc(6)Hex(4)Fuc(0)NeuAc(0)	9	0.195	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2028.7404	N	90	HexNAc(6)Hex(5)Fuc(0)NeuAc(0)	11	0.133	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/2190.7932	N	90	HexNAc(6)Hex(6)Fuc(0)NeuAc(0)	12	0.0716	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1987.7138	N	90	HexNAc(5)Hex(6)Fuc(0)NeuAc(0)	14	0.0451	Undecorated
EQSTLAQMYPLQEIQNLTVK	NGlycan/1298.4760	N	90	HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	16	0.0407	Undecorated
EQSTLAQMYPLQEIQNLTVK	Glu->pyro-Glu/- 18.0106; NGlycan/1702.5813	E,N	75,90	HexNAc(2)Hex(8)Fuc(0)NeuAc(0)	1	0.114	High Mannose
EQSTLAQMYPLQEIQNLTVK LQLQALQQNGSSVLSEDK	Glu->pyro-Glu/- 18.0106; NGlycan/2569.9046	E,N	75,90	HexNAc(5)Hex(6)Fuc(0)NeuAc(2)	14	0.0386	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/2059.7349	N	103	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	5.19	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2059.7349	Q,N	101,103	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	0.0322	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2350.8304	N	103	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	3.91	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/1897.6821	N	103	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	1.79	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1897.6821	Q,N	98,103	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.0119	Sialyfucoylated

LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1897.6821	Q,N	101,103	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.0118	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1897.6821	Q,N	102,103	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.00758	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2367.8457	N	103	HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	4	1.47	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2367.8457	Q,N	101,103	HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	4	0.0161	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2205.7928	N	103	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	5	0.595	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2424.8671	N	103	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	7	0.0542	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2715.9625	N	103	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	8	0.0695	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2262.8143	N	103	HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	9	0.0782	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/3007.0580	N	103	HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	13	0.0252	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2221.7878	Q,N	101,103	HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	15	0.153	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/1694.6027	N	103	HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	18	0.0109	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2716.9829	N	103	HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	22	0.0145	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2391.8569	N	103	HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	30	0.0604	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2408.8722	N	103	HexNAc(5)Hex(5)Fuc(2)NeuAc(1)	42	0.0177	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2408.8722	Q,N	101,103	HexNAc(5)Hex(5)Fuc(2)NeuAc(1)	42	0.00984	Sialyfucoylated
LQLQALQQNGSSVLSEDK	NGlycan/2586.9200	N	103	HexNAc(5)Hex(7)Fuc(1)NeuAc(1)	45	0.00619	Sialyfucoylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2075.7298	Q,N	98,103	HexNAc(4)Hex(6)Fuc(0)NeuAc(1)	1	2.24	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/1954.7036	N	103	HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	2	0.0519	Sialylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2116.7564	Q,N	101,103	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	0.0274	Sialylated

LQLQALQQNGSSVLSEDK	NGlycan/2116.7564	N	103	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	0.174	Sialylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1913.6770	Q,N	98,103	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.44	Sialylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1913.6770	Q,N	102,103	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.429	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/1913.6770	N	103	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.0206	Sialylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1913.6770	Q,N	101,103	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.0057	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/2407.8518	N	103	HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	5	0.0204	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/1751.6242	N	103	HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	9	0.317	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/2204.7724	N	103	HexNAc(4)Hex(5)Fuc(0)NeuAc(2)	10	0.469	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/2278.8092	N	103	HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	13	0.0777	Sialylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/2278.8092	Q,N	98,103	HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	13	0.0382	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/2366.8253	N	103	HexNAc(4)Hex(6)Fuc(0)NeuAc(2)	16	0.012	Sialylated
LQLQALQQNGSSVLSEDK	NGlycan/1444.5339	N	103	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	5.02	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1444.5339	Q,N	101,103	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.0661	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1444.5339	Q,N	98,103	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.0119	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1606.5867	N	103	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	3.26	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1606.5867	Q,N	101,103	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	0.0406	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1768.6395	N	103	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	2.65	Fucosylated

LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1768.6395	Q,N	101,103	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	1.14	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1768.6395	Q,N	96,103	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	0.0296	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2076.7502	N	103	HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	4	2.04	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2133.7717	N	103	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	0.0536	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1914.6974	N	103	HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	6	0.497	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1914.6974	Q,N	101,103	HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	6	0.0298	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1914.6974	Q,N	98,103	HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	6	0.0209	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2222.8082	N	103	HexNAc(4)Hex(6)Fuc(3)NeuAc(0)	8	0.348	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1809.6661	N	103	HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	9	0.187	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1752.6446	N	103	HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	10	0.334	Fucosylated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1752.6446	Q,N	102,103	HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	10	0.141	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2279.8296	N	103	HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	11	0.0819	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1647.6132	N	103	HexNAc(5)Hex(3)Fuc(1)NeuAc(0)	12	0.212	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1971.7189	N	103	HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	13	0.121	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2117.7768	N	103	HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	15	0.0524	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1241.4545	N	103	HexNAc(3)Hex(3)Fuc(1)NeuAc(0)	17	0.0676	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2060.7553	N	103	HexNAc(4)Hex(5)Fuc(3)NeuAc(0)	18	0.114	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2263.8347	N	103	HexNAc(5)Hex(5)Fuc(3)NeuAc(0)	20	0.0615	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1403.5073	N	103	HexNAc(3)Hex(4)Fuc(1)NeuAc(0)	25	0.0489	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2425.8875	N	103	HexNAc(5)Hex(6)Fuc(3)NeuAc(0)	26	0.0459	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/1955.7240	N	103	HexNAc(5)Hex(4)Fuc(2)NeuAc(0)	31	0.0246	Fucosylated
LQLQALQQNGSSVLSEDK	NGlycan/2345.8878	N	103	HexNAc(7)Hex(3)Fuc(3)NeuAc(0)	37	0.000733	Fucosylated

LQLQALQQNGSSVLSEDK	NGlycan/1460.5288	N	103	HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	1	4.24	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1460.5288	Q,N	101,103	HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	1	2.61	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1460.5288	Q,N	98,103	HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	1	2.61	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1460.5288	Q,N	102,103	HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	1	2.55	Undecorated
LQLQALQQNGSSVLSEDK	NGlycan/1622.5816	N	103	HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	2	3.29	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1622.5816	Q,N	98,103	HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	2	1.93	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1622.5816	Q,N	101,103	HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	2	1.88	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1784.6344	Q,N	101,103	HexNAc(4)Hex(6)Fuc(0)NeuAc(0)	3	1.75	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1663.6082	Q,N	102,103	HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	4	0.142	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1663.6082	Q,N	98,103	HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	4	0.142	Undecorated
LQLQALQQNGSSVLSEDK	NGlycan/1663.6082	N	103	HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	4	0.094	Undecorated
LQLQALQQNGSSVLSEDK	NGlycan/1825.6610	N	103	HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	5	0.319	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1825.6610	Q,N	101,103	HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	5	0.113	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1825.6610	Q,N	102,103	HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	5	0.113	Undecorated

LQLQALQQNGSSVLSEDK	NGlycan/1501.5553	N	103	HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	10	0.135	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1257.4494	Q,N	98,103	HexNAc(3)Hex(4)Fuc(0)NeuAc(0)	15	0.0438	Undecorated
LQLQALQQNGSSVLSEDK	NGlycan/1298.4760	N	103	HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	17	0.036	Undecorated
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1298.4760	Q,N	101,103	HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	17	0.0216	Undecorated
LQLQALQQNGSSVLSEDK	NGlycan/2188.7398	N	103	HexNAc(2)Hex(11)Fuc(0)NeuAc(0)	3	0.00855	High Mannose
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1216.4229	Q,N	101,103	HexNAc(2)Hex(5)Fuc(0)NeuAc(0)	2	0.00151	High Mannose
LQLQALQQNGSSVLSEDK	Deamidated/0.984 0; NGlycan/1216.4229	Q,N	102,103	HexNAc(2)Hex(5)Fuc(0)NeuAc(0)	2	0.00151	High Mannose
LQLQALQQNGSSVLSEDKSK	NGlycan/2059.7349	N	103	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	0.175	Sialyfucoylated
LQLQALQQNGSSVLSEDKSK	NGlycan/2350.8304	N	103	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	0.143	Sialyfucoylated
LQLQALQQNGSSVLSEDKSK	NGlycan/1897.6821	N	103	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.0589	Sialyfucoylated
LQLQALQQNGSSVLSEDKSK	NGlycan/2205.7928	N	103	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	5	0.0213	Sialyfucoylated
LQLQALQQNGSSVLSEDKSK	NGlycan/1444.5339	N	103	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.21	Fucosylated
LQLQALQQNGSSVLSEDKSK	Deamidated/0.984 0; NGlycan/1444.5339	Q,N	102,103	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.00489	Fucosylated
LQLQALQQNGSSVLSEDKSK	NGlycan/1606.5867	N	103	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	0.153	Fucosylated
LQLQALQQNGSSVLSEDKSK	NGlycan/1768.6395	N	103	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	0.113	Fucosylated
LQLQALQQNGSSVLSEDKSK	NGlycan/1914.6974	N	103	HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	6	0.0273	Fucosylated
LQLQALQQNGSSVLSEDKSK	NGlycan/1752.6446	N	103	HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	10	0.0213	Fucosylated
FFVSVGLPNMTQGFWENS MLTDPGNVQK	NGlycan/3195.1628	N	322	HexNAc(6)Hex(5)Fuc(4)NeuAc(2)	19	0.0995	Sialyfucoylated
FFVSVGLPNMTQGFWENS MLTDPGNVQK	NGlycan/2538.9352	N	322	HexNAc(5)Hex(4)Fuc(4)NeuAc(1)	20	0.0957	Sialyfucoylated

FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2831.0259	N	322	HexNAc(7)Hex(6)Fuc(1)NeuAc(1)	21	0.0878	Sialyfucoylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2498.9039 ; Deamidated/0.984 0	N,Q	322,325	HexNAc(6)Hex(7)Fuc(1)NeuAc(0)	21	0.00397	Fucosylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2742.0146	N	322	HexNAc(6)Hex(4)Fuc(4)NeuAc(1)	23	0.0757	Sialyfucoylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2085.7870	N	322	HexNAc(5)Hex(3)Fuc(4)NeuAc(0)	28	0.0423	Fucosylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2288.8663	N	322	HexNAc(6)Hex(3)Fuc(4)NeuAc(0)	29	0.0334	Fucosylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2450.9192	N	322	HexNAc(6)Hex(4)Fuc(4)NeuAc(0)	30	0.0261	Fucosylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2904.0674	N	322	HexNAc(6)Hex(5)Fuc(4)NeuAc(1)	31	0.0576	Sialyfucoylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2612.9720	N	322	HexNAc(6)Hex(5)Fuc(4)NeuAc(0)	34	0.0181	Fucosylated
FFVSVGLPNMTQGFWENS MLTDPGNAVQK	NGlycan/2246.8194	N	322	HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	39	0.0279	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2059.7349	N	432	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	0.253	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2350.8304	N	432	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	0.223	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/1897.6821	N	432	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.0157	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2367.8457	N	432	HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	4	0.0248	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2205.7928	N	432	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	5	0.0243	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2789.9993	N	432	HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	6	0.141	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2424.8671	N	432	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	7	0.287	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2715.9625	N	432	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	8	0.282	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2262.8143	N	432	HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	9	0.193	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2100.7615	N	432	HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	10	0.0708	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/3081.0947	N	432	HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	11	0.206	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2732.9779	N	432	HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	12	0.189	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/3007.0580	N	432	HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	13	0.181	Sialyfucoylated

SIGLLSPDFQEDNETEINFLK	NGlycan/2553.9097	N	432	HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	14	0.062	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/3024.0733	N	432	HexNAc(5)Hex(7)Fuc(2)NeuAc(2)	16	0.139	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2627.9465	N	432	HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	26	0.0683	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/3372.1902	N	432	HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	27	0.0659	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2919.0419	N	432	HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	28	0.0651	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2936.0572	N	432	HexNAc(6)Hex(7)Fuc(2)NeuAc(1)	32	0.049	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2570.9250	N	432	HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	33	0.0436	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/3227.1527	N	432	HexNAc(6)Hex(7)Fuc(2)NeuAc(2)	35	0.0339	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2862.0205	N	432	HexNAc(5)Hex(6)Fuc(2)NeuAc(2)	40	0.0234	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/3210.1373	N	432	HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	41	0.021	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2879.0358	N	432	HexNAc(5)Hex(7)Fuc(3)NeuAc(1)	43	0.0164	Sialyfucoylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2075.7298 ; Deamidated/0.984 0	N,N	432,437	HexNAc(4)Hex(6)Fuc(0)NeuAc(1)	1	0.00657	Sialylated
SIGLLSPDFQEDNETEINFLK	Deamidated/0.984 0; NGlycan/2116.7564	Q,N	429,432	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	0.00458	Sialylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2278.8092 ; Deamidated/0.984 0	N,N	432,437	HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	13	0.0152	Sialylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2366.8253 ; Deamidated/0.984 0	N,N	432,437	HexNAc(4)Hex(6)Fuc(0)NeuAc(2)	16	0.0239	Sialylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2643.9414	N	432	HexNAc(6)Hex(7)Fuc(0)NeuAc(1)	18	0.0218	Sialylated
SIGLLSPDFQEDNETEINFLK	Deamidated/0.984 0; NGlycan/2440.8620	Q,N	429,432	HexNAc(5)Hex(7)Fuc(0)NeuAc(1)	20	0.00462	Sialylated
SIGLLSPDFQEDNETEINFLK	NGlycan/1606.5867	N	432	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	0.0226	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/1768.6395	N	432	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	0.0729	Fucosylated

SIGLLSPDFQEDNETEINFLK	NGlycan/2076.7502	N	432	HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	4	0.113	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2133.7717	N	432	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	0.164	Fucosylated
SIGLLSPDFQEDNETEINFLK	Deamidated/0.984 0; NGlycan/2133.7717	Q,N	429,432	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	0.0263	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/1809.6661	N	432	HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	9	0.0374	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2279.8296	N	432	HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	11	0.0759	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/1971.7189	N	432	HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	13	0.0896	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2441.8824	N	432	HexNAc(5)Hex(7)Fuc(2)NeuAc(0)	14	0.173	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2117.7768	N	432	HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	15	0.0209	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2498.9039	N	432	HexNAc(6)Hex(7)Fuc(1)NeuAc(0)	21	0.0564	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2336.8511	N	432	HexNAc(6)Hex(6)Fuc(1)NeuAc(0)	22	0.0541	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2174.7983	N	432	HexNAc(6)Hex(5)Fuc(1)NeuAc(0)	32	0.0231	Fucosylated
SIGLLSPDFQEDNETEINFLK	NGlycan/2295.8245	N	432	HexNAc(5)Hex(7)Fuc(1)NeuAc(0)	33	0.0235	Fucosylated
SIGLLSPDFQEDNETEINFLK	Deamidated/0.984 0; NGlycan/2149.7666	Q,N	429,432	HexNAc(5)Hex(7)Fuc(0)NeuAc(0)	18	0.0203	Undecorated
SIGLLSPDFQEDNETEINFLK	NGlycan/1784.6344 ; Deamidated/0.984 0	N,N	432,437	HexNAc(4)Hex(6)Fuc(0)NeuAc(0)	19	0.0164	Undecorated
SIGLLSPDFQEDNETEINFLK	Deamidated/0.984 0; NGlycan/1825.6610	Q,N	429,432	HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	20	0.0116	Undecorated
SIGLLSPDFQEDNETEINFLK	Deamidated/0.984 0; NGlycan/2352.8460	Q,N	429,432	HexNAc(6)Hex(7)Fuc(0)NeuAc(0)	21	0.00396	Undecorated
CDISNSTEAGQK	NGlycan/2059.7349	N	546	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	0.463	Sialyfucosylated
CDISNSTEAGQK	NGlycan/2350.8304	N	546	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	0.126	Sialyfucosylated
CDISNSTEAGQK	NGlycan/1897.6821	N	546	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.691	Sialyfucosylated
CDISNSTEAGQK	NGlycan/2789.9993	N	546	HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	6	0.0137	Sialyfucosylated
CDISNSTEAGQK	NGlycan/2424.8671	N	546	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	7	0.0957	Sialyfucosylated

CDISNSTEAGQK	NGlycan/2715.9625	N	546	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	8	0.0422	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2262.8143	N	546	HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	9	0.17	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2100.7615	N	546	HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	10	0.208	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2553.9097	N	546	HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	14	0.0278	Sialyfucoylated
CDISNSTEAGQK	NGlycan/1694.6027	N	546	HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	18	0.112	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2465.8937	N	546	HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	24	0.0382	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2627.9465	N	546	HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	26	0.0266	Sialyfucoylated
CDISNSTEAGQK	NGlycan/1856.6556	N	546	HexNAc(3)Hex(5)Fuc(1)NeuAc(1)	29	0.0629	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2303.8409	N	546	HexNAc(6)Hex(4)Fuc(1)NeuAc(1)	34	0.0426	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2351.8508	N	546	HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	38	0.0296	Sialyfucoylated
CDISNSTEAGQK	NGlycan/2756.9891	N	546	HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	44	0.0135	Sialyfucoylated
CDISNSTEAGQK	NGlycan/1954.7036	N	546	HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	2	0.289	Sialylated
CDISNSTEAGQK	NGlycan/2407.8518	N	546	HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	5	0.144	Sialylated
CDISNSTEAGQK	NGlycan/2319.8358	N	546	HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	6	0.00909	Sialylated
CDISNSTEAGQK	NGlycan/2772.9840	N	546	HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	7	0.0157	Sialylated
CDISNSTEAGQK	NGlycan/1751.6242	N	546	HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	9	0.0345	Sialylated
CDISNSTEAGQK	NGlycan/2610.9312	N	546	HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	11	0.0285	Sialylated
CDISNSTEAGQK	NGlycan/1548.5448	N	546	HexNAc(3)Hex(4)Fuc(0)NeuAc(1)	19	0.0206	Sialylated
CDISNSTEAGQK	NGlycan/1444.5339	N	546	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.418	Fucosylated
CDISNSTEAGQK	NGlycan/1606.5867	N	546	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	1.06	Fucosylated
CDISNSTEAGQK	NGlycan/1768.6395	N	546	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	0.251	Fucosylated
CDISNSTEAGQK	NGlycan/2133.7717	N	546	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	0.0436	Fucosylated
CDISNSTEAGQK	NGlycan/1809.6661	N	546	HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	9	0.343	Fucosylated
CDISNSTEAGQK	NGlycan/1752.6446	N	546	HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	10	0.0208	Fucosylated
CDISNSTEAGQK	NGlycan/2279.8296	N	546	HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	11	0.0738	Fucosylated
CDISNSTEAGQK	NGlycan/1647.6132	N	546	HexNAc(5)Hex(3)Fuc(1)NeuAc(0)	12	0.141	Fucosylated
CDISNSTEAGQK	NGlycan/1971.7189	N	546	HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	13	0.182	Fucosylated
CDISNSTEAGQK	NGlycan/1241.4545	N	546	HexNAc(3)Hex(3)Fuc(1)NeuAc(0)	17	0.129	Fucosylated
CDISNSTEAGQK	NGlycan/1565.5601	N	546	HexNAc(3)Hex(5)Fuc(1)NeuAc(0)	19	0.0644	Fucosylated
CDISNSTEAGQK	NGlycan/1850.6926	N	546	HexNAc(6)Hex(3)Fuc(1)NeuAc(0)	23	0.0496	Fucosylated
CDISNSTEAGQK	NGlycan/1403.5073	N	546	HexNAc(3)Hex(4)Fuc(1)NeuAc(0)	25	0.0427	Fucosylated

CDISNSTEAGQK	NGlycan/2012.7454	N	546	HexNAc(6)Hex(4)Fuc(1)NeuAc(0)	27	0.0454	Fucosylated
CDISNSTEAGQK	NGlycan/2174.7983	N	546	HexNAc(6)Hex(5)Fuc(1)NeuAc(0)	32	0.024	Fucosylated
CDISNSTEAGQK	NGlycan/1898.7025	N	546	HexNAc(4)Hex(4)Fuc(3)NeuAc(0)	35	0.00866	Fucosylated
CDISNSTEAGQK	NGlycan/1727.6130	N	546	HexNAc(3)Hex(6)Fuc(1)NeuAc(0)	36	0.00721	Fucosylated
CDISNSTEAGQK	NGlycan/1216.4229	N	546	HexNAc(2)Hex(5)Fuc(0)NeuAc(0)	2	0.0209	High Mannose
NVSDIIPR	NGlycan/2059.7349	N	690	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1	1.75	Sialyfuosylated
NVSDIIPR	NGlycan/2350.8304	N	690	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	2	0.52	Sialyfuosylated
NVSDIIPR	NGlycan/1897.6821	N	690	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	3	0.61	Sialyfuosylated
NVSDIIPR	NGlycan/2205.7928	N	690	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	5	0.0253	Sialyfuosylated
NVSDIIPR	NGlycan/2424.8671	N	690	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	7	0.202	Sialyfuosylated
NVSDIIPR	NGlycan/2715.9625	N	690	HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	8	0.179	Sialyfuosylated
NVSDIIPR	NGlycan/2262.8143	N	690	HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	9	0.0468	Sialyfuosylated
NVSDIIPR	NGlycan/2100.7615	N	690	HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	10	0.0335	Sialyfuosylated
NVSDIIPR	NGlycan/2732.9779	N	690	HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	12	0.134	Sialyfuosylated
NVSDIIPR	NGlycan/3007.0580	N	690	HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	13	0.0296	Sialyfuosylated
NVSDIIPR	NGlycan/3024.0733	N	690	HexNAc(5)Hex(7)Fuc(2)NeuAc(2)	16	0.0197	Sialyfuosylated
NVSDIIPR	NGlycan/2716.9829	N	690	HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	22	0.0858	Sialyfuosylated
NVSDIIPR	NGlycan/2351.8508	N	690	HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	38	0.00339	Sialyfuosylated
NVSDIIPR	NGlycan/2116.7564	N	690	HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	3	0.0204	Sialylated
NVSDIIPR	NGlycan/1913.6770	N	690	HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	4	0.756	Sialylated
NVSDIIPR	NGlycan/1751.6242	N	690	HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	9	0.193	Sialylated
NVSDIIPR	NGlycan/2204.7724	N	690	HexNAc(4)Hex(5)Fuc(0)NeuAc(2)	10	0.101	Sialylated
NVSDIIPR	NGlycan/2278.8092	N	690	HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	13	0.0671	Sialylated
NVSDIIPR	NGlycan/2569.9046	N	690	HexNAc(5)Hex(6)Fuc(0)NeuAc(2)	14	0.0422	Sialylated
NVSDIIPR	NGlycan/1444.5339	N	690	HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	1	0.0189	Fucosylated
NVSDIIPR	NGlycan/1606.5867	N	690	HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	2	0.322	Fucosylated
NVSDIIPR	NGlycan/1768.6395	N	690	HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	3	0.817	Fucosylated
NVSDIIPR	NGlycan/2133.7717	N	690	HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	5	0.113	Fucosylated
NVSDIIPR	NGlycan/1914.6974	N	690	HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	6	0.0243	Fucosylated
NVSDIIPR	NGlycan/1809.6661	N	690	HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	9	0.0703	Fucosylated
NVSDIIPR	NGlycan/1971.7189	N	690	HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	13	0.0882	Fucosylated

NVSDIIPR	NGlycan/2441.8824	N	690	HexNAc(5)Hex(7)Fuc(2)NeuAc(0)	14	0.0876	Fucosylated
NVSDIIPR	NGlycan/1930.6923	N	690	HexNAc(4)Hex(6)Fuc(1)NeuAc(0)	16	0.0822	Fucosylated
NVSDIIPR	NGlycan/1622.5816	N	690	HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	8	0.2	Undecorated
NVSDIIPR	NGlycan/1460.5288	N	690	HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	13	0.0604	Undecorated

Table S2.2 Site-specific occupancy of recombinant ACE2 derived from HEK293 cells

Glycans	Peptide Sequence	Glycosite	Glycan Subtype
HexNAc(5)Hex(5)Fuc(3)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Fucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Fucosylated
HexNAc(9)Hex(9)Fuc(1)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Fucosylated
HexNAc(6)Hex(5)Fuc(4)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(5)Hex(4)Fuc(4)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(5)Hex(5)Fuc(2)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(5)Hex(5)Fuc(4)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(7)Hex(7)Fuc(1)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(7)Hex(7)Fuc(2)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(7)Hex(8)Fuc(1)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(5)Hex(7)Fuc(3)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialyucosylated
HexNAc(7)Hex(8)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(4)Hex(7)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(4)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(5)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated

HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(5)Fuc(0)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(7)Fuc(0)NeuAc(1)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(6)Fuc(0)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(3)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(7)Fuc(0)NeuAc(2)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(7)Hex(7)Fuc(0)NeuAc(3)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Sialylated
HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(6)Hex(3)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(6)Hex(4)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(7)Hex(3)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(4)Hex(7)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(7)Hex(4)Fuc(0)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	53	Undecorated
HexNAc(5)Hex(4)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(3)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(3)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(4)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(6)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(5)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(5)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(5)Fuc(3)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(3)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(4)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(7)Hex(7)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(6)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(7)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(3)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated

HexNAc(4)Hex(3)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(4)Fuc(2)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(3)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(6)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(4)Hex(7)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(7)Hex(6)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Fucosylated
HexNAc(2)Hex(8)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	High Mannose
HexNAc(2)Hex(9)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	High Mannose
HexNAc(6)Hex(4)Fuc(4)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(5)Hex(4)Fuc(3)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(6)Hex(5)Fuc(3)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(6)Hex(4)Fuc(3)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(7)Hex(6)Fuc(3)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(6)Hex(7)Fuc(3)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(6)Hex(5)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(6)Hex(4)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(5)Hex(5)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucosylated

HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.LQLQALQQNGSSVLSEDK.S	90	Sialylated
HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(4)Hex(6)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(5)Hex(7)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(7)Hex(5)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(3)Hex(4)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(4)Hex(6)Fuc(0)NeuAc(2)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(4)Hex(7)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(7)Hex(4)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(7)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(3)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated
HexNAc(6)Hex(3)Fuc(0)NeuAc(1)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Sialylated

HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(4)Hex(6)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(6)Hex(3)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(6)Hex(4)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(6)Hex(5)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(6)Hex(6)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(3)Hex(3)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(7)Hex(6)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(7)Hex(3)Fuc(0)NeuAc(0)	K.EQSTLAQMYPLQEIQNLTVK.L	90	Undecorated
HexNAc(7)Hex(4)Fuc(4)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(5)Fuc(4)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(3)Fuc(4)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(7)Hex(3)Fuc(4)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(4)Fuc(3)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(5)Fuc(3)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(3)Fuc(3)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(6)Fuc(3)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(6)Hex(5)Fuc(3)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(7)Hex(4)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(4)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(3)Hex(4)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(6)Hex(6)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated

HexNAc(5)Hex(7)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(3)Hex(5)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(6)Hex(7)Fuc(2)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(3)Hex(4)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(3)Hex(3)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(3)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(3)Hex(6)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(6)Hex(4)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(4)Hex(6)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(7)Hex(4)Fuc(1)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Fucosylated
HexNAc(5)Hex(4)Fuc(3)NeuAc(0)	K.LQLQALQQNGSSVLSEDKSK.R	103	Fucosylated
HexNAc(2)Hex(11)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	High Mannose
HexNAc(2)Hex(5)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	High Mannose
HexNAc(2)Hex(7)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	High Mannose
HexNAc(2)Hex(10)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	High Mannose
HexNAc(2)Hex(9)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	High Mannose
HexNAc(2)Hex(8)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	High Mannose
HexNAc(4)Hex(4)Fuc(3)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated
HexNAc(5)Hex(4)Fuc(3)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated
HexNAc(5)Hex(5)Fuc(3)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated
HexNAc(4)Hex(6)Fuc(3)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated
HexNAc(5)Hex(7)Fuc(3)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucosylated

HexNAc(9)Hex(9)Fuc(2)NeuAc(4)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(4)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(6)Hex(4)Fuc(2)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(6)Fuc(1)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(7)Hex(4)Fuc(1)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialyfucoylated
HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(7)Hex(4)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated

HexNAc(3)Hex(4)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(3)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(4)Hex(6)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(6)Hex(3)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(5)Hex(7)Fuc(0)NeuAc(1)	K.LQLQALQQNGSSVLSEDK.S	103	Sialylated
HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(6)Hex(7)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(7)Hex(8)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(6)Hex(3)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(6)Hex(4)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(6)Hex(5)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(4)Hex(6)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(3)Hex(4)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(3)Hex(5)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(6)Hex(6)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(5)Hex(6)Fuc(0)NeuAc(0)	K.LQLQALQQNGSSVLSEDK.S	103	Undecorated
HexNAc(7)Hex(6)Fuc(5)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(5)Hex(3)Fuc(4)NeuAc(0)	K.EAEKFFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(5)Hex(4)Fuc(4)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(6)Hex(4)Fuc(4)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(6)Hex(6)Fuc(3)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(6)Hex(3)Fuc(3)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(7)Hex(3)Fuc(2)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(3)Hex(4)Fuc(2)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated

HexNAc(5)Hex(3)Fuc(2)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(7)Hex(4)Fuc(2)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(4)Hex(7)Fuc(1)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(4)Hex(6)Fuc(1)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(3)Hex(5)Fuc(1)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Fucosylated
HexNAc(2)Hex(6)Fuc(0)NeuAc(0)	K.EAEKFFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	High Mannose
HexNAc(2)Hex(11)Fuc(0)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	High Mannose
HexNAc(2)Hex(4)Fuc(0)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	High Mannose
HexNAc(7)Hex(5)Fuc(5)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(4)Fuc(5)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(4)Fuc(4)NeuAc(1)	K.EAEKFFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(5)Hex(4)Fuc(4)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(5)Fuc(4)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(5)Fuc(4)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(7)Hex(5)Fuc(4)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(5)Hex(4)Fuc(3)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(4)Fuc(3)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(9)Hex(9)Fuc(3)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(6)Fuc(3)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(7)Hex(5)Fuc(3)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(4)Hex(7)Fuc(3)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(4)Fuc(2)NeuAc(1)	K.EAEKFFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(3)Fuc(2)NeuAc(1)	K.EAEKFFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(4)Hex(4)Fuc(2)NeuAc(1)	K.EAEKFFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(5)Fuc(2)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(8)Hex(8)Fuc(2)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(5)Hex(7)Fuc(2)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(7)Hex(5)Fuc(2)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(4)Hex(7)Fuc(2)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated

HexNAc(3)Hex(4)Fuc(2)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(7)Hex(4)Fuc(2)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(3)Hex(5)Fuc(1)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(6)Hex(3)Fuc(1)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(7)Hex(5)Fuc(1)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(7)Hex(4)Fuc(1)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(3)Hex(6)Fuc(1)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialyucosylated
HexNAc(4)Hex(6)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(3)Hex(4)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(4)Hex(6)Fuc(0)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(4)Hex(7)Fuc(0)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(7)Hex(5)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(6)Hex(3)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(3)Hex(5)Fuc(0)NeuAc(1)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Sialylated
HexNAc(6)Hex(3)Fuc(0)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Undecorated
HexNAc(7)Hex(6)Fuc(0)NeuAc(0)	K.FFVSVGLPNMTQGFWENSMLTDPGNVQK.A	322	Undecorated
HexNAc(4)Hex(5)Fuc(3)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(4)Fuc(3)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(6)Fuc(3)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(5)Fuc(3)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(5)Fuc(2)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(7)Fuc(2)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated

HexNAc(6)Hex(5)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(5)Hex(3)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(6)Hex(4)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Fucosylated
HexNAc(6)Hex(7)Fuc(4)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(4)Fuc(3)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(7)Fuc(2)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(7)Fuc(2)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(6)Fuc(2)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(7)Fuc(2)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(7)Fuc(2)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyucosylated

HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(4)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialyfucoylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	K.HLSIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(7)Hex(6)Fuc(0)NeuAc(2)	K.HLSIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(6)Hex(7)Fuc(0)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(7)Hex(7)Fuc(0)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(7)Hex(7)Fuc(0)NeuAc(3)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(2)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(6)Hex(7)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Undecorated
HexNAc(6)Hex(4)Fuc(0)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Undecorated
HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Undecorated
HexNAc(4)Hex(6)Fuc(0)NeuAc(0)	K.SIGLLSPDFQEDNETEINFLK.Q	432	Undecorated
HexNAc(4)Hex(4)Fuc(3)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(5)Fuc(3)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated

HexNAc(4)Hex(5)Fuc(3)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(6)Fuc(3)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(4)Fuc(3)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(4)Hex(4)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(4)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(3)Hex(4)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(3)Fuc(2)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(3)Hex(4)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(3)Hex(5)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(3)Hex(3)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(6)Hex(4)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(5)Hex(3)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(6)Hex(3)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(3)Hex(6)Fuc(1)NeuAc(0)	K.CDISNSTEAGQK.L	546	Fucosylated
HexNAc(2)Hex(8)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	High Mannose
HexNAc(2)Hex(5)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	High Mannose
HexNAc(2)Hex(7)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	High Mannose
HexNAc(2)Hex(6)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	High Mannose
HexNAc(2)Hex(11)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	High Mannose
HexNAc(5)Hex(4)Fuc(4)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucosylated

HexNAc(6)Hex(4)Fuc(4)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(3)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(3)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(4)Hex(6)Fuc(2)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(4)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(3)Hex(5)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(3)Hex(6)Fuc(1)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialyfucoylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialylated

HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(3)Hex(4)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(7)Hex(5)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(6)Hex(3)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(7)Hex(4)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(3)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(3)Hex(5)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(7)Hex(6)Fuc(0)NeuAc(1)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(7)Hex(7)Fuc(0)NeuAc(3)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(7)Hex(6)Fuc(0)NeuAc(2)	K.CDISNSTEAGQK.L	546	Sialylated
HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(6)Hex(3)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(3)Hex(5)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(6)Hex(5)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(6)Hex(4)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(3)Hex(4)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(3)Hex(6)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(6)Hex(6)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(7)Hex(5)Fuc(0)NeuAc(0)	K.CDISNSTEAGQK.L	546	Undecorated
HexNAc(7)Hex(5)Fuc(4)NeuAc(0)	K.FNHEAEDLFYQSSLASWNYNTNITEENVQNMNNAGDK.W	690	Fucosylated
HexNAc(5)Hex(5)Fuc(4)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated

HexNAc(5)Hex(7)Fuc(2)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(3)Hex(4)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(4)Hex(3)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(3)Hex(3)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(5)Hex(7)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(0)	K.NVSDIIPR.T	690	Fucosylated
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(7)Fuc(2)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(5)Fuc(2)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	K.NVSDIIPR.T	690	Sialyfucosylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated

HexNAc(4)Hex(4)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(4)Hex(5)Fuc(0)NeuAc(2)	K.NVSDIIPR.T	690	Sialylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(5)Hex(4)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(2)	K.NVSDIIPR.T	690	Sialylated
HexNAc(5)Hex(5)Fuc(0)NeuAc(2)	K.NVSDIIPR.T	690	Sialylated
HexNAc(6)Hex(6)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(5)Hex(6)Fuc(0)NeuAc(2)	K.NVSDIIPR.T	690	Sialylated
HexNAc(6)Hex(7)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(6)Hex(4)Fuc(0)NeuAc(1)	K.NVSDIIPR.T	690	Sialylated
HexNAc(6)Hex(5)Fuc(0)NeuAc(2)	K.NVSDIIPR.T	690	Sialylated
HexNAc(3)Hex(4)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(4)Hex(5)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(4)Hex(4)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(4)Hex(3)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(5)Hex(5)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(6)Hex(7)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(3)Hex(3)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(5)Hex(6)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(5)Hex(4)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(5)Hex(3)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated
HexNAc(6)Hex(6)Fuc(0)NeuAc(0)	K.NVSDIIPR.T	690	Undecorated

Table S2.3 HepG2 N-Glycome Profiles

Treatment	Mass (exp.)	Glycan Subtype	Hex	HexNAc	Fuc	NeuAc	RT	Relative Abundance
Control	3171.1077	Sialyfucoylated	6	5	2	3	33.79	17.09
Control	3025.05215	Sialyfucoylated	6	5	1	3	33.435	11.45

Control	2368.82772	Sialyfucoylated	5	4	1	2	27.272	9.59
Control	2880.01648	Sialyfucoylated	6	5	2	2	28.099	6.53
Control	3536.23075	Sialyfucoylated	7	6	2	3	32.88	4.60
Control	2733.95668	Sialyfucoylated	6	5	1	2	26.792	4.43
Control	2222.77629	Sialylated	5	4	0	2	26.406	3.73
Control	1720.58797	High Mannose	8	2	0	0	16.527	2.81
Control	1558.53507	High Mannose	7	2	0	0	16.535	2.30
Control	2878.99649	Sialylated	6	5	0	3	30.249	2.25
Control	3099.08627	Sialyfucoylated	7	6	1	2	28.928	2.25
Control	1882.63764	High Mannose	9	2	0	0	14.876	2.06
Control	1396.48489	High Mannose	6	2	0	0	16.534	1.90
Control	1234.43361	High Mannose	5	2	0	0	15.055	1.73
Control	3245.14147	Sialyfucoylated	7	6	2	2	28.276	1.68
Control	2077.74411	Sialyfucoylated	5	4	1	1	25.034	1.53
Control	3901.36548	Sialyfucoylated	8	7	2	3	32.562	1.52
Control	4047.42108	Sialyfucoylated	8	7	3	3	32.674	1.26
Control	2807.99108	Sialyfucoylated	7	6	1	1	25.401	1.20
Control	4412.54442	Sialyfucoylated	9	8	3	3	31.832	1.11
Control	2587.9302	Sialylated	6	5	0	2	26.246	1.11
Control	4558.60993	Sialyfucoylated	9	8	4	3	32.39	1.10
Control	4265.47873	Sialylated	9	8	0	4	33.248	0.78
Control	2514.90407	Sialyfucoylated	5	4	2	2	28.523	0.75
Control	3026.06317	Sialyfucoylated	6	5	3	2	28.563	0.73
Control	3682.2928	Sialyfucoylated	7	6	3	3	33.002	0.73
Control	3683.31196	Sialylated	9	8	0	2	33.296	0.64
Control	4048.42004	Sialyfucoylated	8	7	5	2	32.699	0.59
Control	3756.32838	Sialyfucoylated	8	7	3	2	28.07	0.57
Control	3391.20449	Sialyfucoylated	7	6	3	2	28.027	0.53
Control	3101.09265	Undecorated	9	8	0	0	28.415	0.49
Control	2881.00488	Sialyfucoylated	6	5	4	1	27.692	0.42
Control	2571.90946	Sialyfucoylated	5	5	1	2	26.295	0.42

Control	2734.96851	Sialyfucoylated	6	5	3	1	33.651	0.39
Control	1931.68258	Sialylated	5	4	0	1	23.753	0.39
Control	2954.05527	Sialyfucoylated	7	6	2	1	26.097	0.38
Control	2044.68308	High Mannose	10	2	0	0	17.203	0.38
Control	1728.60378	Sialylated	5	3	0	1	21.681	0.38
Control	3431.13406	Sialyfucoylated	6	7	1	3	31.982	0.36
Control	2223.80247	Sialyfucoylated	5	4	2	1	25.502	0.35
Control	1072.37862	High Mannose	4	2	0	0	16.521	0.31
Control	1890.65345	Sialylated	6	3	0	1	22.723	0.30
Control	2953.05105	Sialylated	7	6	0	2	26.139	0.29
Control	3757.30547	Sialyfucoylated	8	7	5	1	32.822	0.27
Control	2589.91795	Fucosylated	6	5	4	0	28.085	0.23
Control	2409.86881	Sialyfucoylated	4	5	1	2	27.005	0.23
Control	2661.94723	Sialylated	7	6	0	1	25.696	0.23
Control	4339.52848	Sialyfucoylated	8	7	5	3	33.275	0.20
Control	2296.80193	Sialylated	6	5	0	1	25.705	0.20
Control	3319.13534	Sialyfucoylated	8	7	2	1	28.781	0.19
Control	4119.41958	Sialyfucoylated	7	6	4	4	35.184	0.19
Control	4485.67268	Sialyfucoylated	8	7	6	3	38.828	0.19
Control	3828.35852	Sialyfucoylated	7	6	4	3	34.004	0.18
Control	1462.5422	Fucosylated	3	4	1	0	20.95	0.18
Control	3100.10068	Sialyfucoylated	7	6	3	1	25.852	0.17
Control	4015.44804	Sialyfucoylated	6	7	5	3	31.983	0.17
Control	2206.75588	High Mannose	11	2	0	0	18.813	0.16
Control	2692.9092	Sialyfucoylated	7	4	1	2	38.675	0.15
Control	2572.96644	Sialyfucoylated	5	5	3	1	29.548	0.15
Control	910.32508	High Mannose	3	2	0	0	17.598	0.15
Control	3390.16773	Sialyfucoylated	7	6	1	3	33.259	0.14
Control	3448.21784	Sialyfucoylated	7	7	2	2	33.233	0.14
Control	3244.1124	Sialylated	7	6	0	3	26.78	0.14
Control	2442.8558	Sialyfucoylated	6	5	1	1	25.67	0.13

Control	5067.7955	Sialyfucoylated	8	7	6	5	36.386	0.13
Control	3830.32217	Sialyfucoylated	9	8	3	1	30.32	0.12
Control	2458.8865	Sialylated	7	5	0	1	37.857	0.12
Control	3522.23455	Sialyfucoylated	8	8	2	1	32.887	0.12
Control	2524.90341	Sialyfucoylated	4	7	1	1	32.428	0.12
Control	1712.61544	Sialyfucoylated	4	3	1	1	25.716	0.12
Control	2614.92315	Fucosylated	4	6	5	0	33.298	0.11
Control	3611.26524	Sialyfucoylated	8	7	4	1	38.783	0.11
Control	2790.97739	Sialylated	6	6	0	2	32.444	0.11
Control	2425.82883	Sialylated	5	5	0	2	35.078	0.10
Control	2921.07879	Sialyfucoylated	5	6	2	2	35.268	0.10
Control	4046.42392	Sialyfucoylated	8	7	1	4	35.085	0.10
Control	3975.39019	Sialyfucoylated	9	8	2	2	31.281	0.10
Control	3318.16418	Sialylated	8	7	0	2	30.164	0.09
Control	2882.10681	Fucosylated	8	7	1	0	30.816	0.09
Control	3027.05103	Sialylated	8	7	0	1	28.578	0.09
Control	2588.89877	Sialyfucoylated	6	5	2	1	36.527	0.09
Control	2384.83165	Sialylated	6	4	0	2	27.744	0.08
Control	2906.09018	Sialyfucoylated	4	6	5	1	38.744	0.08
Control	1275.44533	Undecorated	4	3	0	0	35.01	0.08
Control	3083.10218	Sialyfucoylated	6	6	2	2	26.23	0.07
Control	3593.25052	Sialyfucoylated	7	7	1	3	31.955	0.07
Control	3539.23755	Fucosylated	9	8	3	0	27.747	0.06
Control	2939.01167	Undecorated	8	8	0	0	36.393	0.06
Control	2645.92309	Sialyfucoylated	6	6	1	1	31.998	0.06
Control	2240.85387	Fucosylated	6	4	3	0	37.912	0.06
Control	2305.78103	Sialyfucoylated	3	6	2	1	37.219	0.06
Control	2834.03916	Fucosylated	5	7	4	0	38.018	0.06
Control	2838.97667	Sialyfucoylated	7	4	2	2	37.941	0.05
Control	2513.90677	Sialylated	5	4	0	3	35.88	0.04
Control	2955.06296	Fucosylated	7	6	4	0	28.231	0.04

Control	2368.82969	High Mannose	12	2	0	0	29.213	0.04
Control	2809.03324	Fucosylated	7	6	3	0	37.715	0.04
Control	2297.81088	Fucosylated	6	5	2	0	29.713	0.03
2F-Fucose	2222.77505	Sialylated	5	4	0	2	25.676	19.62
2F-Fucose	2878.99658	Sialylated	6	5	0	3	26.037	16.78
2F-Fucose	2953.03182	Sialylated	7	6	0	2	27.705	6.79
2F-Fucose	2587.90476	Sialylated	6	5	0	2	26.533	6.78
2F-Fucose	1720.58802	High Mannose	8	2	0	0	16.566	3.77
2F-Fucose	2661.94302	Sialylated	7	6	0	1	25.267	3.75
2F-Fucose	1931.67633	Sialylated	5	4	0	1	23.079	3.42
2F-Fucose	3244.11253	Sialylated	7	6	0	3	28.846	2.92
2F-Fucose	1558.53822	High Mannose	7	2	0	0	15.041	2.90
2F-Fucose	2296.80792	Sialylated	6	5	0	1	24.025	2.77
2F-Fucose	1882.647	High Mannose	9	2	0	0	14.123	2.75
2F-Fucose	1234.43339	High Mannose	5	2	0	0	15.19	2.45
2F-Fucose	1396.48898	High Mannose	6	2	0	0	16.568	2.42
2F-Fucose	2368.82765	Sialyfucoylated	5	4	1	2	27.779	1.86
2F-Fucose	3318.16515	Sialylated	8	7	0	2	29.098	1.74
2F-Fucose	3609.25396	Sialylated	8	7	0	3	33.503	1.32
2F-Fucose	3683.29518	Sialylated	9	8	0	2	29.583	1.25
2F-Fucose	3974.38315	Sialylated	9	8	0	3	33.786	1.13
2F-Fucose	2879.99453	Sialyfucoylated	6	5	2	2	30.617	0.89
2F-Fucose	1728.60884	Sialylated	5	3	0	1	21.694	0.88
2F-Fucose	1890.65652	Sialylated	6	3	0	1	22.729	0.78
2F-Fucose	2588.91014	Sialyfucoylated	6	5	2	1	24.98	0.72
2F-Fucose	3025.04442	Sialyfucoylated	6	5	1	3	36.657	0.71
2F-Fucose	2733.93209	Sialyfucoylated	6	5	1	2	28.397	0.62
2F-Fucose	2954.02543	Sialyfucoylated	7	6	2	1	28.189	0.60
2F-Fucose	3099.09833	Sialyfucoylated	7	6	1	2	28.855	0.58
2F-Fucose	2044.68014	High Mannose	10	2	0	0	17.216	0.53
2F-Fucose	3026.06469	Sialyfucoylated	6	5	3	2	26.219	0.52

2F-Fucose	1072.38139	High Mannose	4	2	0	0	16.577	0.50
2F-Fucose	2589.95618	Fucosylated	6	5	4	0	26.636	0.50
2F-Fucose	2881.99063	Fucosylated	8	7	1	0	25.486	0.46
2F-Fucose	3685.28046	Fucosylated	9	8	4	0	29.58	0.46
2F-Fucose	3027.06498	Sialylated	8	7	0	1	26.755	0.44
2F-Fucose	2425.85821	Sialylated	5	5	0	2	25.171	0.43
2F-Fucose	3245.11443	Sialyfucoylated	7	6	2	2	33.647	0.36
2F-Fucose	3976.36614	Sialyfucoylated	9	8	4	1	34.047	0.32
2F-Fucose	910.3236	High Mannose	3	2	0	0	17.626	0.29
2F-Fucose	2790.9762	Sialylated	6	6	0	2	25.61	0.28
2F-Fucose	1566.56591	Sialylated	4	3	0	1	23.026	0.27
2F-Fucose	2206.74945	High Mannose	11	2	0	0	18.823	0.26
2F-Fucose	2660.97388	Sialyfucoylated	5	4	3	2	21.898	0.24
2F-Fucose	3682.2912	Sialyfucoylated	7	6	3	3	29.574	0.24
2F-Fucose	2077.73078	Sialyfucoylated	5	4	1	1	25.728	0.24
2F-Fucose	2368.80395	High Mannose	12	2	0	0	27.192	0.21
2F-Fucose	4048.40764	Sialyfucoylated	8	7	5	2	29.825	0.21
2F-Fucose	2735.92624	Undecorated	8	7	0	0	28.945	0.18
2F-Fucose	3611.25006	Sialyfucoylated	8	7	4	1	35.844	0.18
2F-Fucose	4413.4887	Sialyfucoylated	9	8	5	2	31.14	0.18
2F-Fucose	2264.80701	Sialyfucoylated	4	5	2	1	25.913	0.17
2F-Fucose	2881.00685	Sialyfucoylated	6	5	4	1	33.124	0.17
2F-Fucose	2734.96526	Sialyfucoylated	6	5	3	1	27.874	0.14
2F-Fucose	2061.72161	Sialyfucoylated	4	4	2	1	37.573	0.14
2F-Fucose	3684.29203	Sialyfucoylated	9	8	2	1	31.638	0.14
2F-Fucose	3320.11472	Fucosylated	8	7	4	0	35.419	0.12
2F-Fucose	3101.14769	Fucosylated	7	6	5	0	39.219	0.12
2F-Fucose	4339.50078	Sialyfucoylated	8	7	5	3	33.904	0.10
2F-Fucose	3246.11681	Sialyfucoylated	7	6	4	1	35.726	0.10
2F-Fucose	2426.89513	Sialyfucoylated	5	5	2	1	35.333	0.10
2F-Fucose	1802.6326	Undecorated	6	4	0	0	33.999	0.10

2F-Fucose	2808.03381	Sialyfucoylated	7	6	1	1	26.387	0.10
2F-Fucose	2499.9437	Sialylated	6	6	0	1	38.209	0.10
2F-Fucose	3392.17516	Sialylated	9	8	0	1	33.655	0.09
2F-Fucose	2297.83217	Fucosylated	6	5	2	0	35.682	0.08
2F-Fucose	1316.48659	Undecorated	3	4	0	0	18.27	0.07
2F-Fucose	2719.93553	Fucosylated	7	7	1	0	37.565	0.07
2F-Fucose	2662.94794	Fucosylated	7	6	2	0	26.188	0.07
2F-Fucose	2304.85609	Sialylated	3	6	0	2	26.475	0.07
2F-Fucose	2994.00598	Sialylated	6	7	0	2	30.169	0.06
2F-Fucose	4705.61901	Sialyfucoylated	9	8	7	2	35.79	0.06
2F-Fucose	2078.76857	Fucosylated	5	4	3	0	25.122	0.06
2F-Fucose	3319.13729	Sialyfucoylated	8	7	2	1	30.648	0.05
2F-Fucose	2546.8532	Sialylated	7	4	0	2	26.248	0.05
2F-Fucose	3903.47421	Sialyfucoylated	8	7	6	1	37.479	0.05
2F-Fucose	2458.91249	Sialylated	7	5	0	1	26.117	0.04
2F-Fucose	2385.89524	Sialyfucoylated	6	4	2	1	35.948	0.04
2F-Fucose	2571.87902	Sialyfucoylated	5	5	1	2	26.287	0.04
Kifunensine	1882.63654	High Mannose	9	2	0	0	18.42	30.04
Kifunensine	1720.58961	High Mannose	8	2	0	0	18.552	23.46
Kifunensine	1558.53554	High Mannose	7	2	0	0	18.571	15.68
Kifunensine	1396.48489	High Mannose	6	2	0	0	18.571	8.29
Kifunensine	1234.43341	High Mannose	5	2	0	0	21.976	5.19
Kifunensine	2044.69275	High Mannose	10	2	0	0	18.422	2.46
Kifunensine	2368.82775	Sialyfucoylated	5	4	1	2	28.266	1.89
Kifunensine	2036.71566	Sialyfucoylated	6	3	1	1	25.084	1.64
Kifunensine	1072.37197	High Mannose	4	2	0	0	18.571	1.52
Kifunensine	1890.65922	Sialylated	6	3	0	1	25.073	0.98
Kifunensine	3318.19664	Sialylated	8	7	0	2	31.922	0.88
Kifunensine	910.32426	High Mannose	3	2	0	0	18.574	0.79
Kifunensine	2206.73848	High Mannose	11	2	0	0	20.126	0.70
Kifunensine	1874.66193	Sialyfucoylated	5	3	1	1	24.953	0.57

Kifunensine	3521.25843	Sialylated	8	8	0	2	37.869	0.48
Kifunensine	2401.83446	Sialyfucoylated	7	4	1	1	35.177	0.41
Kifunensine	3084.13601	Sialyfucoylated	6	6	4	1	35.582	0.40
Kifunensine	1728.60681	Sialylated	5	3	0	1	25.007	0.38
Kifunensine	2368.77118	High Mannose	12	2	0	0	33.846	0.36
Kifunensine	2662.96639	Fucosylated	7	6	2	0	35.27	0.34
Kifunensine	2906.16005	Sialyfucoylated	4	6	5	1	28.748	0.31
Kifunensine	2540.93262	Sialylated	5	7	0	1	39.016	0.30
Kifunensine	2255.75782	Sialylated	7	4	0	1	31.505	0.29
Kifunensine	2880.00539	Sialyfucoylated	6	5	2	2	28.562	0.28
Kifunensine	2222.77518	Sialylated	5	4	0	2	27.397	0.27
Kifunensine	2735.99792	Undecorated	8	7	0	0	38.275	0.26
Kifunensine	2217.87719	Fucosylated	3	7	2	0	37.524	0.20
Kifunensine	1915.70408	Sialyfucoylated	4	4	1	1	36.118	0.18
Kifunensine	2459.89419	Fucosylated	7	5	2	0	36.872	0.17
Kifunensine	2160.76488	Fucosylated	3	6	3	0	34.838	0.16
Kifunensine	2524.98709	Sialyfucoylated	4	7	1	1	38.967	0.15
Kifunensine	2450.86613	Sialyfucoylated	3	6	1	2	25.093	0.13
Kifunensine	4179.44551	Sialyfucoylated	9	9	4	1	35.099	0.12
Kifunensine	2134.72875	Sialylated	5	5	0	1	33.378	0.11
Kifunensine	2369.84004	Sialyfucoylated	5	4	3	1	32.443	0.11
Kifunensine	2379.84386	Fucosylated	4	7	2	0	32.923	0.11
Kifunensine	1891.64516	Fucosylated	6	3	2	0	21.918	0.11
Kifunensine	2734.02197	Sialyfucoylated	6	5	1	2	30.822	0.09
Kifunensine	2547.8558	Sialyfucoylated	7	4	2	1	35.576	0.09
Kifunensine	2530.82552	Sialyfucoylated	6	4	1	2	32.544	0.07
Kifunensine	2573.92736	Undecorated	7	7	0	0	30.666	0.05
3-F-Sia	1720.586	High Mannose	8	2	0	0	17.353	6.88
3-F-Sia	2516.9077	Fucosylated	7	6	1	0	25.941	5.50
3-F-Sia	1558.53538	High Mannose	7	2	0	0	17.352	5.39
3-F-Sia	2077.74286	Sialyfucoylated	5	4	1	1	28.389	5.18

3-F-Sia	2368.83364	Sialyfucoylated	5	4	1	2	30.593	4.85
3-F-Sia	1882.63829	High Mannose	9	2	0	0	17.219	4.26
3-F-Sia	1396.48476	High Mannose	6	2	0	0	17.443	4.25
3-F-Sia	1234.4309	High Mannose	5	2	0	0	21.106	3.94
3-F-Sia	1786.64097	Fucosylated	5	4	1	0	25.258	3.48
3-F-Sia	2370.83605	Undecorated	7	6	0	0	25.019	3.23
3-F-Sia	2808.00462	Sialyfucoylated	7	6	1	1	27.074	2.99
3-F-Sia	2954.04937	Sialyfucoylated	7	6	2	1	26.213	2.36
3-F-Sia	2297.81326	Fucosylated	6	5	2	0	25.431	2.09
3-F-Sia	2882.02771	Fucosylated	8	7	1	0	26.098	2.05
3-F-Sia	2588.92525	Sialyfucoylated	6	5	2	1	26.396	1.94
3-F-Sia	2223.79455	Sialyfucoylated	5	4	2	1	25.675	1.90
3-F-Sia	1932.67907	Fucosylated	5	4	2	0	24.976	1.87
3-F-Sia	2442.86866	Sialyfucoylated	6	5	1	1	26.62	1.74
3-F-Sia	3028.09659	Fucosylated	8	7	2	0	28.437	1.71
3-F-Sia	1640.58591	Undecorated	5	4	0	0	35.013	1.71
3-F-Sia	2880.04654	Sialyfucoylated	6	5	2	2	38.959	1.70
3-F-Sia	2809.00778	Fucosylated	7	6	3	0	23.875	1.63
3-F-Sia	3099.08639	Sialyfucoylated	7	6	1	2	29.617	1.59
3-F-Sia	2005.71767	Undecorated	6	5	0	0	23.093	1.45
3-F-Sia	3100.09993	Sialyfucoylated	7	6	3	1	29.537	1.43
3-F-Sia	3245.14437	Sialyfucoylated	7	6	2	2	30.516	1.40
3-F-Sia	2662.94933	Fucosylated	7	6	2	0	25.466	1.38
3-F-Sia	3247.15	Fucosylated	9	8	1	0	26.713	1.32
3-F-Sia	2222.7747	Sialylated	5	4	0	2	29.578	1.08
3-F-Sia	2044.69093	High Mannose	10	2	0	0	20.016	0.92
3-F-Sia	3319.16483	Sialyfucoylated	8	7	2	1	26.542	0.87
3-F-Sia	1072.37752	High Mannose	4	2	0	0	17.437	0.79
3-F-Sia	1599.55907	Undecorated	6	3	0	0	19.552	0.70
3-F-Sia	3539.26307	Fucosylated	9	8	3	0	25.305	0.65
3-F-Sia	3465.23302	Sialyfucoylated	8	7	3	1	26.339	0.57

3-F-Sia	1712.62035	Sialyfucoylated	4	3	1	1	28.383	0.51
3-F-Sia	3537.25022	Sialyfucoylated	7	6	4	2	28.274	0.49
3-F-Sia	2661.94328	Sialylated	7	6	0	1	25.8	0.49
3-F-Sia	3246.15536	Sialyfucoylated	7	6	4	1	25.689	0.49
3-F-Sia	3976.40533	Sialyfucoylated	9	8	4	1	26.483	0.48
3-F-Sia	2078.75633	Fucosylated	5	4	3	0	21.163	0.48
3-F-Sia	2206.73801	High Mannose	11	2	0	0	20.112	0.48
3-F-Sia	3391.19318	Sialyfucoylated	7	6	3	2	28.495	0.47
3-F-Sia	3464.2053	Sialyfucoylated	8	7	1	2	30.585	0.47
3-F-Sia	2587.90396	Sialylated	6	5	0	2	29.646	0.45
3-F-Sia	1275.4572	Undecorated	4	3	0	0	22.809	0.45
3-F-Sia	1421.51798	Fucosylated	4	3	1	0	21.978	0.41
3-F-Sia	2734.97407	Sialyfucoylated	6	5	3	1	26.182	0.41
3-F-Sia	1462.54135	Fucosylated	3	4	1	0	20.631	0.40
3-F-Sia	1728.60241	Sialylated	5	3	0	1	23.495	0.39
3-F-Sia	1437.5089	Undecorated	5	3	0	0	19.535	0.37
3-F-Sia	2719.98218	Fucosylated	7	7	1	0	20.622	0.35
3-F-Sia	2384.85617	Sialylated	6	4	0	2	28.267	0.35
3-F-Sia	1890.65742	Sialylated	6	3	0	1	24.28	0.33
3-F-Sia	3393.20671	Fucosylated	9	8	2	0	26.06	0.31
3-F-Sia	2093.73431	Sialylated	6	4	0	1	25.876	0.29
3-F-Sia	3903.37858	Sialyfucoylated	8	7	6	1	28.421	0.29
3-F-Sia	4123.46536	Fucosylated	9	8	7	0	26.959	0.28
3-F-Sia	2645.93503	Sialyfucoylated	6	6	1	1	32.336	0.27
3-F-Sia	2953.02357	Sialylated	7	6	0	2	28.453	0.25
3-F-Sia	1259.46769	Fucosylated	3	3	1	0	24.147	0.25
3-F-Sia	3685.28777	Fucosylated	9	8	4	0	37.541	0.24
3-F-Sia	2589.9414	Fucosylated	6	5	4	0	22.834	0.20
3-F-Sia	910.32577	High Mannose	3	2	0	0	21.013	0.20
3-F-Sia	3174.14215	Fucosylated	8	7	3	0	25.208	0.20
3-F-Sia	3831.35379	Fucosylated	9	8	5	0	26.683	0.19

3-F-Sia	1915.67879	Sialyfucoylated	4	4	1	1	26.613	0.19
3-F-Sia	1665.6196	Fucosylated	3	5	1	0	21.993	0.19
3-F-Sia	3026.06328	Sialyfucoylated	6	5	3	2	29.012	0.19
3-F-Sia	1868.70072	Fucosylated	3	6	1	0	21.446	0.18
3-F-Sia	1583.57208	Fucosylated	5	3	1	0	20.471	0.18
3-F-Sia	3466.24153	Fucosylated	8	7	5	0	26.414	0.18
3-F-Sia	3536.22715	Sialyfucoylated	7	6	2	3	31.09	0.17
3-F-Sia	2354.8569	Fucosylated	6	6	1	0	19.958	0.17
3-F-Sia	2630.99329	Fucosylated	5	6	4	0	36.596	0.15
3-F-Sia	2955.02886	Fucosylated	7	6	4	0	28.724	0.15
3-F-Sia	1989.71387	Fucosylated	5	5	1	0	23.973	0.14
3-F-Sia	3757.31879	Sialyfucoylated	8	7	5	1	25.677	0.14
3-F-Sia	2368.80585	High Mannose	12	2	0	0	21.669	0.14
3-F-Sia	2134.71958	Sialylated	5	5	0	1	37.628	0.13
3-F-Sia	1931.69133	Sialylated	5	4	0	1	29.624	0.13
3-F-Sia	2264.81909	Sialyfucoylated	4	5	2	1	25.283	0.12
3-F-Sia	3684.2584	Sialyfucoylated	9	8	2	1	33.301	0.10
3-F-Sia	1745.61187	Fucosylated	6	3	1	0	18.619	0.10
3-F-Sia	3612.25987	Fucosylated	8	7	6	0	28.504	0.09
3-F-Sia	2897.03326	Sialyfucoylated	7	5	3	1	36.039	0.09
3-F-Sia	2151.77901	Fucosylated	6	5	1	0	26.534	0.09
3-F-Sia	2094.74813	Fucosylated	6	4	2	0	36.069	0.08
3-F-Sia	2547.90255	Sialyfucoylated	7	4	2	1	35.668	0.08
3-F-Sia	4414.56797	Sialyfucoylated	9	8	7	1	28.255	0.07
3-F-Sia	2822.93785	Sialyfucoylated	6	4	3	2	37.122	0.07
3-F-Sia	2087.80714	Undecorated	4	7	0	0	32.639	0.06
3-F-Sia	4996.74472	Sialyfucoylated	9	8	7	3	36.324	0.06
3-F-Sia	2572.88395	Sialyfucoylated	5	5	3	1	26.79	0.06
3-F-Sia	2280.8109	Sialyfucoylated	5	5	1	1	23.018	0.06
3-F-Sia	2426.86719	Sialyfucoylated	5	5	2	1	26.83	0.06
3-F-Sia	2978.06525	Sialyfucoylated	5	7	1	2	34.581	0.05

3-F-Sia	2239.79153	Sialyfucoylated	6	4	1	1	25.629	0.05
3-F-Sia	2646.93468	Fucosylated	6	6	3	0	34.239	0.05
3-F-Sia	2790.96657	Sialylated	6	6	0	2	33.314	0.04
3-F-Sia	3010.05383	Sialyfucoylated	5	5	4	2	37.026	0.04
3-F-Sia	2369.84416	Sialyfucoylated	5	4	3	1	18.8	0.03
3-F-Sia	2864.97364	Sialylated	7	7	0	1	36.659	0.03
3-F-Sia	2530.87403	Sialyfucoylated	6	4	1	2	37.909	0.02
3-F-Sia	2499.88748	Sialylated	6	6	0	1	28.377	0.02

Table S2.4 Site-specific occupancy of recombinant Spike protein RBD derived from HEK293 cells

Glycans	Peptide Sequence	Glycosite	Glycan Subtype
HexNAc(8)Hex(9)Fuc(3)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	331	Fucosylated
HexNAc(9)Hex(9)Fuc(3)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	331	Fucosylated
HexNAc(9)Hex(9)Fuc(6)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	331	Fucosylated
HexNAc(9)Hex(9)Fuc(2)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(7)Hex(6)Fuc(5)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(5)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(1)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(4)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(7)Hex(7)Fuc(4)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(7)Hex(6)Fuc(5)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(0)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	331	Sialylated
HexNAc(8)Hex(9)Fuc(7)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(7)Hex(6)Fuc(4)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(7)Hex(8)Fuc(6)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(7)Hex(7)Fuc(2)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	331	Sialyfucoylated
HexNAc(8)Hex(9)Fuc(2)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(3)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(6)Hex(7)Fuc(4)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated

HexNAc(9)Hex(9)Fuc(2)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(6)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(8)Hex(9)Fuc(3)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(9)Hex(9)Fuc(4)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(3)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(7)Hex(7)Fuc(2)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(6)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(2)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(8)Hex(8)Fuc(6)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(7)Hex(6)Fuc(2)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(7)Hex(8)Fuc(6)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(4)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(9)Hex(9)Fuc(1)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(9)Fuc(5)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(3)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(9)Hex(9)Fuc(2)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(9)Fuc(2)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(8)Hex(8)Fuc(5)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(7)Hex(8)Fuc(5)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(9)Fuc(6)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(7)Hex(6)Fuc(4)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(2)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(9)Hex(9)Fuc(5)NeuAc(0)	R.FPNITNLCPFGEVFNATR.F	343	Fucosylated
HexNAc(8)Hex(9)Fuc(3)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(2)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(5)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(6)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(7)Hex(6)Fuc(3)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(9)Fuc(3)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(7)Hex(8)Fuc(6)NeuAc(3)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(2)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated

HexNAc(8)Hex(9)Fuc(7)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(9)Hex(9)Fuc(3)NeuAc(2)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated
HexNAc(8)Hex(8)Fuc(6)NeuAc(1)	R.FPNITNLCPFGEVFNATR.F	343	Sialyfucoylated

Chapter 3 Systemically Modifying Cell Glycocalyx to Understand Host-*Salmonella* Interaction

Authors:

Ying Sheng^{1, 2}; Yixuan (Axe) Xie¹; Maurice Wong¹; Carlito B Lebrilla*^{1, 2}

¹Department of Chemistry, University of California, Davis, CA

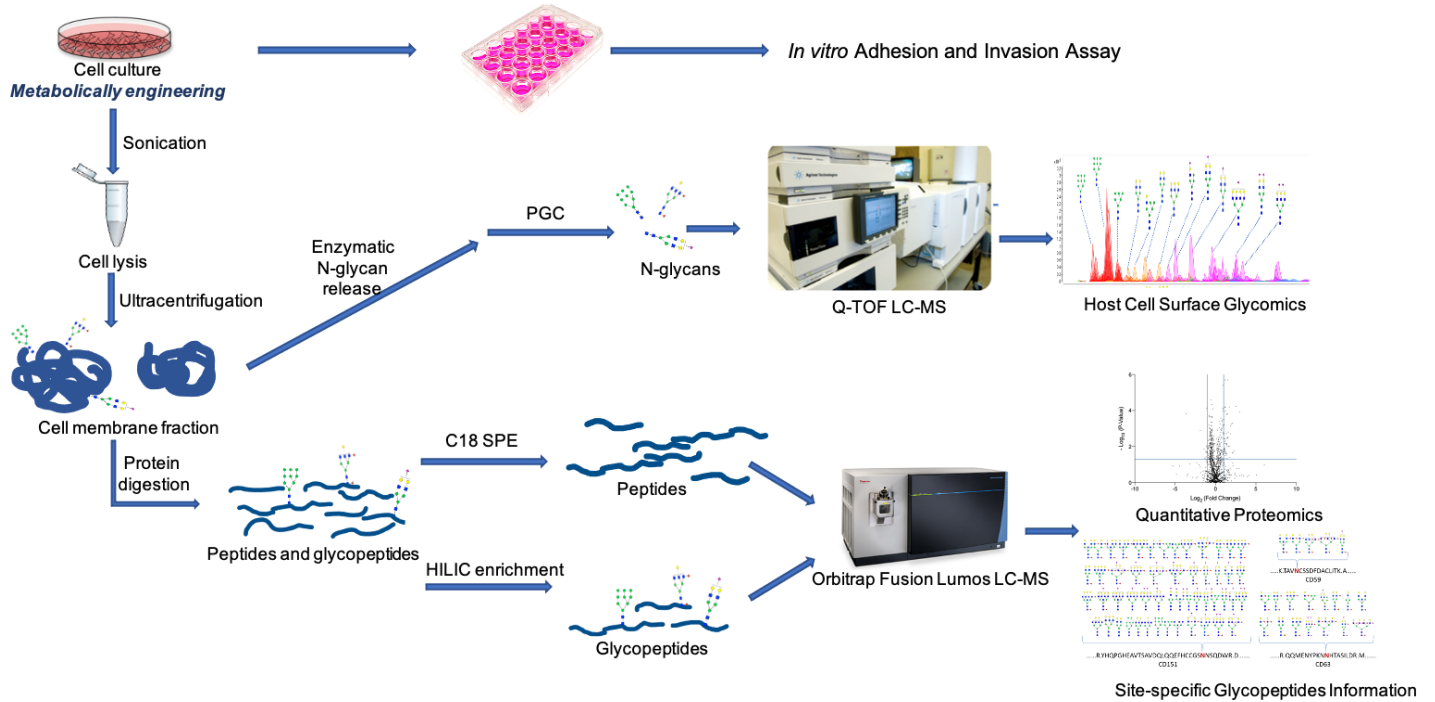
²The Biochemistry, Molecular, Cellular, and Developmental Biology (BMCDB) Graduate Group, University of California, Davis, CA

ABSTRACT

Host-microbe interactions are mediated by protein-carbohydrate binding processes. N-Glycans are oligosaccharides attached to the polypeptide of proteins and are found on the surface of mammalian cells. The pathogenic bacterium, *Salmonella enterica* serotype Typhi (*S. Typhi*), was employed in this study. The host cell surface glycome was manipulated via metabolic engineering. We established a cell-based model that enabled us to perform reliable structure-phenotype correlative experiments and compare the effect of individual N-glycan subtypes in bacterial infection. We created host cell surfaces that were primarily fucosylated, sialylated, undecorated, or oligomannose structures using specific inhibitors. The host cell glycomes were characterized by Q-TOF LC-MS. Adherence assays showed fucosylated N-glycans on the colon cancer cell line HCT116 cell surface significantly increased the number of adhered *S. Typhi*. Furthermore, adherence of *S. Typhi* to HCT116 cells could be blocked by co-incubation with fucose. The results proved that fucose residues on host cells bind with *S. Typhi* during bacterial infection. Proteomic analysis showed that the glycoengineering did not change the abundance of membrane proteins, which indicated that host protein expression did not contribute to the increased adherence of *S. Typhi*. Meanwhile, glycoproteomic analysis yielded site-specific N-glycosylation information. Glycopeptides were identified and quantified using a standard glycoproteomic workflow. These

results suggested the importance of glycans in host-microbe interactions and provided a novel insight into the significance of host glycome during pathogenesis.

GRAPHICAL ABSTRACT



INTRODUCTION

Host-microbe interactions are often mediated by protein-carbohydrate binding processes(1-3). Thus, it is critical to provide efficient tools to reveal the common mechanisms involved in these interactions. Highly glycosylated epithelial cells within the mammalian gut are the primary boundary separating embedded host tissues from intestinal pathogens. Comprising the outermost layer of all eukaryotic cells, glycans are complicated carbohydrate structures that mediate host interactions with billions of native microorganisms, antibodies, and other host cells(1, 4, 5). N-Glycans are covalently attached to proteins at asparagine (Asn) residues and present on many secreted and membrane-bound glycoproteins at the Asn in Asn-X-Ser/Thr sequons(6). About 70% of proteins contain this sequon, and ~70% of sequons carry an N-glycan(7). N-Glycans can be grouped into subtypes, which include high mannose (HM), nondecorated complex/hybrid (C/H), fucosylated C/H, and sialylated C/H(8). Currently, the biological functions of some glycan compounds in bacterial pathogenesis have been identified(9, 10), but the ability to explore the roles of glycans during bacterial infection is still limited.

L-Fucose, the only levorotatory monosaccharide utilized by mammalian systems, is highly abundant in the stomach and large intestine but usually present at low levels in the small intestine(11). This fact makes fucose a useful location marker for bacteria. For example, the chemotactic behavior of *C. jejuni* toward L-fucose has been discovered in previous studies (12, 13). Further evidence in literature suggests that fucosylated glycans regulate bacterial intestinal colonization with other mechanisms(9, 14, 15). Besides the potential use as a carbon and energy source, fucosylated glycans can serve as adherence sites or receptors for pathogens, including *H. pylori*, enterotoxigenic *E. coli*, norovirus, and *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*)(9, 13, 16).

The adherence of pathogenic bacteria to cellular targets in mammalian tissues is crucial in the pathogenesis of many infections. Gram-negative bacteria have lectin-like adhesive molecules(*i.e.adhesins*) to interact with glycans on the host cell surface, contributing to bacterial attachment(2). Adhesins are usually carried by a hair-like structure called fimbriae. Pathogenic bacteria possess various fimbriae, and each fimbriae recognizes and binds to certain receptors to facilitate adherence(3). For example, long polar fimbria (LPF) and type 1 fimbria allow bacteria to interact with mannose(17). Type 4 fimbriae binds to the Lewis X (Le^x) blood group antigen(18). Std fimbriae recognizes terminal Fuca1-2 residues(9).

Salmonella is a facultative anaerobic Gram-negative rod-shaped bacterium that colonizes within the intestinal tract of humans and farm animals(19). *Salmonella* contains two species, *S. enterica* and *S. bongori*, and includes more than 2,579 serovars. *Salmonella enterica* serovar Typhi (*S. Typhi*) is a medically important pathogen that infects the intestinal tract and the blood and causes typhoid fever, a systemic disease of humans, estimated to cause 216,000 deaths from 21 million cases (20). However, the effect of the host cell surface glycome on *S. Typhi* infection has not been comprehensively explored. Details are still unclear regarding the host glycan-mediated interaction of *S. Typhi* and host cells.

To compare the effects of different N-glycan subtypes on *S. Typhi* infection within one cell line, this study has established an innovative cell-based model in which the cell surface glycome in a colon cancer cell line, HCT116, can be efficiently manipulated via metabolic pathway engineering. This model enables us to perform reliable structure-phenotype correlative experiments and compare the effects of individual N-glycan subtypes in *S. Typhi* infection within the same host cell line. We determined that the host fucosylated glycans are involved in *S. Typhi* infection combined with adherence and invasion assays. The proteomic analysis confirmed that host plasma membrane proteins do not contribute to the increased adherence of *S. Typhi*. Characterizing the site-specific N-glycosylation of cell membrane glycoproteins provides potential host targets involved in *S. Typhi* infection. We

developed this platform to map both the potential glycans and glycoproteins interacting with pathogen through glycomic and glycoproteomic analyses. This study evaluated associations between glycosylation of host proteins and bacterial adherence. We reported that greater fucosylation of proteins leads to increased severity of infection.

METHODS AND MATERIALS

Cell culture and inhibitors treatment

Human colorectal cancer HCT116 cells were obtained from American Type Culture Collection (ATCC, VA) and grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum and 100 U mL⁻¹ penicillin and streptomycin. Cells were subcultured at 90% confluency and maintained at 37 °C in a humidified incubator with 5% CO₂. At 50% cell confluency, the cells were treated with 100 μM kifunensine, 2-fluoro-L-fucose, 3-fluorinated sialic acid in cell culture media and let to stand in the incubator for 48 hours.

Bacterial culture

Salmonella enterica serotype Typhi (*S. Typhi*) (ATCC 19430) were grown at 37 °C in Luria-Bertani (LB) broth (1% tryptone, 1% sodium chloride, 0.5% yeast extract) with continuous shaking for 14–16 h. The bacterial culture was then inoculated with 2% in a 250-mL Erlenmeyer flask containing 50 ml of LB broth and grew at 37°C with continuous shaking (220 rpm) for 6 hours.

Bacterial Adherence and Invasion Assays

Cell medium was changed 18-24 hours before infection. Caco-2 cells were washed the cells with warm PBS and, to each well, added 0.5 ml of fresh medium supplemented with 16.6 % serum but containing no antibiotics. Prior to adherence and invasion assays, epithelial cells were washed with serum-free, antibiotic-free medium. Cells were infected with *S. Typhi* (10⁸ CFU/well) and allowed to be co-incubated for 60 min at 37°C. The cell number was counted to calculate the MOI (the multiplicity of infection). Following the incubation, non-

adherent bacteria were removed by two washes with PBS. Adherent bacteria resuspended in PBS containing 1% (v/v) Triton X-100. Serial 10-fold dilutions were spread on agar plates to determine the number of cell-associated bacteria per well. In invasion experiments, non-adherent bacteria were washed as described above, and cells were subsequently incubated at room temperature in a medium containing 0.1 mg/ mL gentamicin for 60 min prior to determining the number of cell-associated bacteria per well as described above.

Cell membrane extraction

The procedures for cell membrane extraction were described previously(8, 21, 22). Cells were collected and resuspended in homogenization buffer containing 0.25 M sucrose, 20 mM HEPES-KOH (pH 7.4), and protease inhibitor mixture (1:100; Calbiochem/EMD Chemicals). Cells were lysed on ice with five alternating on and off pulses in 5 and 10-second intervals using a probe sonicator (Qsonica, CT). Nuclear and mitochondrial fractions and cellular debris were pelleted and isolated by centrifugation at $2000 \times g$ for 10 min. The supernatants were then ultra-centrifuged at $200\,000 \times g$ for 45 min at 4 °C to extract the plasma membrane. The pellets of the cell membrane were resuspended in 500 μ L of 0.2 M Na_2CO_3 solution and 500 μ L of water followed by two more ultracentrifugation treatments at $200\,000 \times g$ for 45 min to wash off the endoplasmic reticulum (ER) and cytoplasmic fraction.

Enzymatic release and purification of N-glycans

Extracted cell membrane fractions(22) were suspended with 100 μ L of 100 mM NH_4HCO_3 in 5 mM dithiothreitol and heated in boiling water for 2 minutes to denature the proteins. Solutions with 2 μ L of peptide N-glycosidase F (New England Biolabs, MA) were added to the samples to release the N-glycans from proteins, and the resulting solutions were then incubated in a microwave reactor (CEM Corporation, NC) at 20 watts, 37 °C for 10 min. The samples were further placed in a 37 °C water bath for 18 hours. Ultracentrifugation at $200\,000 \times g$ for 30 min was performed to precipitate the proteins. The supernatant containing N-glycans was

collected and purified using porous graphitic carbon (PGC) on a 96-well SPE plate (Grace, IL). The plate was equilibrated with 80% (v/v) acetonitrile containing 0.1% (v/v) trifluoroacetic acid. The samples were loaded onto the plate and washed with nanopure water. N-Glycans were eluted with a solution of 40% (v/v) acetonitrile containing 0.05% (v/v) trifluoroacetic acid, and the samples were dried *in vacuo* using miVac (SP Scientific, PA) prior to mass spectrometric analysis.

Protein digestion

Details of protein digestion have been described previously(22). Extracted cell membrane proteins were reconstituted in 60 μ L of 8 M urea at room temperature. Dissolved cell membrane proteins were reduced with 2 μ L of 550 mM dithiothreitol, alkylated with 4 μ L of 450 mM iodoacetamide. A 420 μ L of 50 mM ammonium bicarbonate was added to dilute the urea concentration to 1 M and to adjust the pH value. The samples were incubated with 2 μ g trypsin at 37 °C for 18 hr. The resulting peptides were concentrated *in vacuo* using miVac (SP Scientific, PA). Glycopeptides were enriched by solid-phase extraction using iSPE[®]-HILIC cartridges (HILICON, Sweden). The cartridges were conditioned with 0.1% (v/v) trifluoroacetic acid in acetonitrile, followed by 1% (v/v) trifluoroacetic acid and 80% (v/v) acetonitrile in water. The samples were loaded and washed with 1% (v/v) trifluoroacetic acid and 80% (v/v) acetonitrile in water. The enriched glycopeptides were eluted with water containing 0.1% (v/v) trifluoroacetic acid and dried prior to mass spectrometric analysis.

Glycomic analysis with LC-MS/MS

Glycan samples were reconstituted(22) in 10 μ L nanopure water and analyzed using an Agilent 6520 Accurate Mass Q-TOF LC/MS equipped with a PGC nano-chip (Agilent Technologies, CA). The glycan separation was performed at a constant flow rate of 300 nL min⁻¹, and a binary gradient was applied using (A) 0.1% (v/v) formic acid in 3% acetonitrile and (B) 1% (v/v) formic acid in 90% acetonitrile: 0–2 min, 0–0% (B); 2–20 min, 0–16% (B); 20–40 min, 16%–72% (B); 40–42 min, 72–100% (B); 42–52 min, 100–100% (B); 52–54 min, 100–0% (B);

54–65 min, 0–0% (B). MS spectra were collected with a mass range of m/z 600–2000 at a rate of 1.5 s per spectrum in positive ionization mode. The most abundant precursor ions in each MS1 spectrum were subjected to be fragmented through collision-induced dissociation (CID) based on the equation $V_{\text{collision}} = 1.8 \times (m/z) / 100 \text{ V} - 2.4 \text{ V}$.

Glycoproteomic analysis with LC-MS/MS

The enriched glycopeptide(22) samples were reconstituted with 20 μL of water and directly characterized using UltiMate™ WPS-3000RS nanoLC 980 system coupled to the Nanospray Flex ion source of an Orbitrap Fusion Lumos Tribrid Mass Spectrometer system (Thermo Fisher Scientific, MA). The analytes were separated on an Acclaim™ PepMap™ 100 C18 LC Column (3 μm , 0.075 mm x 150 mm, ThermoFisher Scientific). A binary gradient was applied using 0.1% (v/v) formic acid in (A) water and (B) 80% acetonitrile: 0–5 min, 4–4% (B); 5–133 min, 4–32% (B); 133–152 min, 32%–48% (B); 152–155 min, 48–100% (B); 155–170 min, 100–100% (B); 170–171 min, 100–4% (B); 171–180 min, 4–4% (B). The instrument was run in data-dependent mode with 1.8kV spray voltage, 275 °C ion transfer capillary temperature. The acquisition was performed with the full MS scanned from 700 to 2000 in positive ionization mode. Stepped higher-energy C-trap dissociation (HCD) at $30 \pm 10\%$ was applied to obtain tandem MS/MS spectra with m/z values starting from 120.

Glycomic data analysis

Extraction of the compound chromatographs of glycans from cells was obtained via the MassHunter Qualitative Analysis B08 software (Agilent, CA). N-Glycan compositions were identified according to accurate mass using an in-house library constructed based on the knowledge of N-glycan biosynthetic pathways and previously obtained in-house structures. Relative abundances were determined by integrating peak areas for observed glycan masses and normalizing to the summed peak areas of all glycans detected.

Glycoproteomic data analysis

Glycopeptide fragmentation spectra were annotated using Byonic software (Protein Metrics, CA) against the reviewed UniProt homo sapiens (human) protein database. Carbamidomethyl modification at cysteine residues and oxidation at methionine were assigned as the modification. The glycan information acquired from the glycomic analysis was employed for the glycopeptide identification.

RESULTS

Metabolically engineering glycosylation in HCT116 cell line

To determine the specific effects of each N-glycan subtype on *S. Typhi* Infection, we developed an *in vitro* model using monosaccharides and inhibitors to enhance or diminish particular glycan subtypes. The colon cancer cell line HCT116 was selected because it showed complete loss of fucosylation due to a mutation in the GDP-fucose synthetic enzyme, GDP-mannose-4,6-dehydratase (GMDS)(23, 24). The N-glycans extracted from cell surfaces were analyzed with an Agilent 6520 Accurate Mass Q-TOF LC/MS equipped with a PGC micro-fluidic chip. Chromatograms of identified glycan compounds from unmodified and modified HCT116 cells were shown in **Figure 3.1**. Based on the sum of all glycan signals, compositional profiles were generated as pie charts to examine general trends according to glycan subtypes. The N-glycome profile showed mainly sialylation (38% sialylated and 3% sialyfucosylated N-glycans) under native conditions.

Applying fluorinated sugars as glycosyltransferase inhibitors is a common glycoengineering strategy(25, 26). The treatment with sialyltransferase inhibitor, 3-fluorinated sialic acid (3-F-Sia), inhibited sialylation, dramatically increasing the relative abundance of undecorated compounds (orange peaks in chromatogram) from 11% to 54%. GDP-Fuc can also be synthesized from fucose via *salvage* pathway(24); therefore the addition of exogenous fucose increased fucosylated N-glycans to 11% and sialyfucosylated ones (blue peaks in chromatogram) to 39%. With the combination of fucose and sialic acid inhibitor, the N-glycans were converted to fucosylated ones (green peaks in chromatogram) and accounted for 37% of the total relative abundances. The treatment of mannosidase inhibitor kifunensine, which blocked mannose trimming during glycan-biosynthesis(27), increased the relative abundance of oligomannose (red peaks in chromatogram) to 96%. The extracted ion chromatograms (EID) proved that our method efficiently manipulates the cell surface glycome within the selected cell line.

***S. Typhi* binds to fucose residues on host cell surface**

Adherence and invasion assays were employed to determine *S. Typhi* infection capacity with different subtypes of N-glycans dominating the host cell surface. The modified cells were employed in adherence and invasion assays. A previously reported study demonstrated that mannose-binding occurred during host-*Salmonella* cell interactions(28). We therefore expected that invasion of HCT116 would significantly increase after kifunensine treatment. Indeed the kifunensine treatment increased the adherence ratio of *S. Typhi* from 7% to 16 % and increased the invasion ratio from 0.08 ‰ to 0.16 ‰ (**Figure 3. 2.**).

On the other hand, the adherence assay displayed another interesting increase. The treatment of the cells combining fucose and sialyltransferase inhibitor increased the *S. Typhi* adherence greatly. The adherence ratio increased from 7 % to almost 60 % (**Figure 3. 2.**). This observation indicated that fucosylated N-glycans on HCT116 cell surface facilitated *S. Typhi* adherence.

The advantage of the HCT116 cell line for these studies is that fucose becomes the only resource for synthesizing GDP-Fuc. The mutation in GMDS cuts off *de novo* fucosylation pathway in this cell line. Therefore, the abundances of fucosylated glycans can be tuned by varying the amount of exogenous fucose in cell media (**Figure 3. 3. A.**). The graph showed a notable tendency that more fucosylated N-glycans were incorporated on cell surface as the gradient of exogenous fucose increased. The ion abundances of fucosylated glycans increased from a level lower than detection limitation to 3×10^7 . These modified cells were then applied in adherence and invasion assays (**Figure 3. 3. B.**). We did not observe a notable change in the number of invaded bacteria from the treated cells. However, the number of adhered bacteria showed a 100-fold increase with 100 μ M fucose. This remarkable change demonstrated that the adherence of *S. Typhi* climbed with the increased abundance of fucosylated glycans.

Based on these data, we hypothesized that *S. Typhi* needs fucose residues on the host cell surface as receptors to initiate infection. To test this hypothesis, *S. Typhi* were preincubated with fucose before interacting with HCT116

cell culture (**Figure 3. 3. C.**). While the concentration of exogenous fucose increased from 5 to 5000 μM , the adherence ratio decreased from 3.5 % to 1.1 %. A more obvious trend was observed in the invasion assay. There remained with 0.4 % *S. Typhi* invasion without exogenous fucose, but the invasion ratio has decreased to 0.09 % when the fucose concentration is just 5 μM . Furthermore, the invasion almost disappeared with the concentration of exogenous fucose higher than 50 μM . The infection of *S. Typhi* to HCT116 cells could be blocked by preincubation with fucoses further supporting the hypothesis that fucose moieties on host cells surface function as a receptor for *S. Typhi* infection. In this experimental design, glucose was used as a monosaccharide control. It is known that glucose is a preferred carbon source for most bacteria(14, 29). Thus, we suspect that the increased invasion with glucose is caused by faster bacterial proliferation due to the carbon source provided during the preincubation step.

To determine whether there were specificity to the fucose linkage, we used oligosaccharides with various linkages from human milk to preincubate with *S. Typhi*. Human milk oligosaccharides (HMOs) 2'-Fucosyllactose (2'-FL), 3'-Fucosyllactose (3'-FL), and 6'-Fucosyllactose (6'-FL) were ideal reagents because they contained a variety of fucose linkages. Each compound was preincubated with *S. Typhi*, and the binding to the host cells were monitored. We observed a significant decrease in adherence with *S. Typhi* when it was preincubated with 3'-FL/6'-FL (**Figure 3. 3. D.**). When the *S. Typhi* were incubated with 3'-FL/6'-FL before adherence, the numbers of attached *S. Typhi* were only half of the ones preincubated with 2'-FL. The differences indicated the *S. Typhi* strain used in this study preferred $\alpha(1,3)$ and $\alpha(1,6)$ over $\alpha(1,2)$ linked fucose.

Proteomic quantification reveals unchanged expression of most proteins on cell surfaces

To rule out the effect of protein expression in modified cells, proteomic analysis was performed to monitor the abundance of proteins from HCT116 cells with and without glycoengineering. Lengths of peptides selected for quantification were between seven and twenty-three amino acids to eliminate the latent occurrence of miss cleavages. Each protein was represented by its most abundant peptide. Volcano plots represented the expression levels of proteins in HCT116 cells with respect to control and treated cells (**Figure 3.**

4.). Around 1200 proteins were identified, and more than 1000 of them did not show significant differences after glycoengineering (fold change >2 or <0.5, and P-Value ≤ 0.05).

Since we concluded that *S. Typhi* binds to fucose residues on host cell surface, we were particularly interested in the protein expression level when the cells became fucosylated. After comparing the untreated cells and the cells treated with the combination of exogenous fucose and sialyltransferase inhibitor (**Figure 3. 4. C.**), we found 62 out of 1,262 proteins showed significant differential expression after treatment. Only a limited number (4 out of 62) were located on the plasma membrane. It showed that only four plasma membrane proteins (P06733, 043688, P15151, and Q9Y639) changed their expression levels when the cell was treated with the combination of exogenous fucose and sialyltransferase inhibitor. None of these proteins have been reported to contribute to bacterial adherence. The data supported that the expression levels of most proteins on the cell surface were not influenced by manipulating glycosylation. The results further supported the unique importance of host glycosylation during pathogenesis.

Glycoproteomic analysis yields proteins associated with glycans

To obtain the host glycoproteins contributing to *S. Typhi* infection, a glycoproteomic analysis was performed to map the glycosylation sites acquired fucosylation after modification. The glycoproteomic analysis identified 372 glycosites across 2228 glycopeptides in HCT116 cells collected from five different treatments. **Figure 3. 5. A** showed the site-specific occupancy of glycoprotein *ITGA5* integrin alpha-5 in HCT116 cells under the five different treatments described above. We found four glycosites on Integrin alpha-5, and **Figure 3. 5. A** presented all the glycans detected on these glycosites. It displayed a clear pattern of glycosylation in this site-specific N-glycan map. For example, the glycan Hex₅HexNAc₄NeuAc₁ detected on glycosite N¹⁸² was sialylated in the control group. In addition, the treatment of 3-F-Sia inhibited the addition of sialic acid (NeuAc), and the glycan composition became Hex₅HexNAc₄. Compared to the glycan from the control group, providing exogenous fucose in cell culture added one more fucose residue to the compound and generated Hex₅HexNAc₄Fuc₁NeuAc₁. In the

cells treated with fucose and sialyltransferase inhibitor simultaneously, glycans lost sialic acids and gained fucoses in their compositions. Hex₅HexNAc₄Fuc₁ and Hex₄HexNAc₄Fuc₁ were yielded on this glycosite. Lastly, kifunensine treatment inhibited mannose trimming and left only oligomannose Hex₈HexNAc₂ on the glycosite N¹⁸².

This trend is more evident when the glycoforms' abundances are quantified. Before the analysis, glycopeptides were enriched to remove the overwhelming abundances of peptides in the cell culture samples. After the enrichment, the same amounts of glycopeptides in each sample were analyzed via Orbitrap Fusion Lumos LC-MS. The relative abundance here studied the relative abundances of individual glycopeptides on each specific glycosite of a given protein independent of the level of that protein. The abundance of each glycoform was normalized to the total abundance of the corresponding glycopeptides, which differed only in their glycan structure. The heat map in **Figure 3. 5. B.** showed the relative abundance of glycopeptides sharing the peptide sequence STDN¹⁸²FTRI. Treatment with fucose and sialic acid inhibitors simultaneously increased the adherence level of *S. Typhi*. It also led to two glycoforms at site N¹⁸² where both were fucosylated. This pattern was consistent with our previous data based on the glycomic profiles.

The pattern displayed in the glycoprotein *ITGA5* (integrin alpha-5) was not always observed. **Figure 3. 5. C.** showed the relative abundances of glycoforms from *ADAM10* (disintegrin and metalloproteinase domain-containing protein 10). Different treatments did not change the attached N-glycans on glycosites N²⁶⁷, N²⁷⁸, and N⁴³⁹. In particular, all the glycans detected on glycosite N⁴³⁹ were oligomannose despite the cell culture treatment. Additionally, Hex₉HexNAc₂ (Man9) occupied this site more frequently than Hex₈HexNAc₂. We can assume that these glycoforms were not associated with the increased bacterial adherence. We predicted the glycoproteins which obtained fucosylation when bacterial adherence is increased, such as *ITGA5* integrin alpha-5, were the potential targets for the microbial interactions.

DISCUSSION

It is known that glycocalyx is optimally positioned as the primary molecular contacts engaged in cellular encounters with viruses, bacteria, antibodies, toxins, and other host cells. Even though glycosylation is being studied in pathogenesis, the effect of host glycocalyx on pathogenesis has not been comprehensively evaluated with live tissue culture thus far. To systematically study the glycan-mediated interactions between different biological entities, some advanced technologies have been applied to identify bacterial-host glycointeractions(3). Most of the recent technological advances in the field of glycointeractomics are high-throughput screening methods, including lectin microarrays(30) and glycan (and mucin) microarrays(31, 32). These methods allow the simultaneous analysis of binding to hundreds of test glycans, but they are limited to the level of glycans and the *in vitro* environment. In contrast, our study developed a mass spectrometry-based platform to probe glycan-mediated interaction *ex vivo*. In other words, the glycomics and glycoproteomics-based workflow developed in this study not only can characterize the binding glycans but also obtain the engaged glycoprotein targets through glycoproteomics. Another advantage of our method is applying bacteria, viruses, or cells that are still active. Unlike our approach, lectin microarrays usually utilize extracted glycoproteins. The initial glycan interaction is less significant in static situations. So this method is more suitable to mimic an actual biological situation. Our results supported the notion that host cell glycocalyx mediates host-microbe interaction and facilitates the study of glycan-mediated host-microbe interactions.

In the human glycome, fucose is most commonly linked to N-glycans at the core region in an α (1,6) configuration. It is also possible to be attached near the terminal end of glycans in the α (1,2), (1,3/4) configurations(10, 16). Studies by Chessa *et al.* have shown that receptors for α (1-2) fucosylated structures are present in the cecal mucosa of mice and bind to adhesins in *S. Typhimurium* (9). Hao *et al.* have reported that core fucosylation(α 1-6) of intestinal epithelial cells protects against infection of *S. Typhi* strain CMCC (B) 50071 CICC 10871 (15). The bacterial protein expression and binding preference vary from strain to strain, even within the same species. Therefore, we

wanted to explore further the linkage preference of the *S. Typhi* strain employed in this study. The data in **Figure 3. 3. D.** showed 3'-FL and 6'-FL were more attractive to *S. Typhi*, suggesting this strain has a higher affinity towards α (1,3) and (1,6), instead of α (1,2) linked fucose moieties. We preincubated modified cells with various lectins to block certain residues, then conducted adherence and invasion assays (**Figure S3. 1.**). The specificities of lectins are summarized in **Table S 3. 1.** (33). Lectin AAL blocks binding to Fuc α 1-3 and Fuc α 1-6 residues. Indeed, the loss of these receptors decreased the number of adhered *S. Typhi*. With the adherence assay data from preincubation with HMOs, we concluded that the *S. Typhi* strain recognizes and binds Fuc α 1-6/ α 1-3 residues on host cells. We also predicted that the infection would be altered by pretreatment of cells with highly specific exoglycosidase, such as α -1,3/4-fucosidase (data not shown here). However, we were unable to confirm the specificity of these fucosidases via mass spectrometry due to lack of appropriate standards and potential fucose rearrangement in the TOF analyzer(34, 35).

There were 103 glycoproteins identified in the glycocalyx of HCT116 cells. Proteomic analysis revealed four plasma membrane proteins (P06733, O43688, P15151, and Q9Y639) that showed significant differences after the cell surface were mainly fucosylated. Only protein Q9Y639 was identified as glycoprotein, while the rest did not present glycans. Glycoprotein Q9Y639 did not display the same pattern as in *ITGA5*. For example, glycosite N²²⁹ was occupied with oligomannoses Hex₆₋₉HexNAc₂ with no and all treatments. Even though these four proteins changed their expression levels significantly, their alternations did not contribute to the increased adherence level we observed. These findings further corroborated our conclusions regarding the significant role of host glycosylation.

Our platform is a novel addition to the glycointeractomic toolbox because it overcomes the weakness that limits traditional lectin-based analysis. It identifies the glycans and the associating proteins in the host-microbe interaction. It also generates glycan compositions and site-specific occupancies of host glycocalyx. The method is not limited to the infection of *S. Typhi*. It will be valuable in identifying targets of other pathogeneses.

The platform's functionality will be expanded in the future to include other diseases that target the cell glycocalyx. Furthermore, this study revealed host fucose residues function as receptors for *S.Typhi*. Since diet provides exogenous fucose for human cells, this research raises the possibility that diet may influence pathogenic infection state via altering glycosylation.

CONCLUSION

We established a cell-based model that enabled us to perform reliable structure-phenotype correlative experiments and compare the effect of individual N-glycan subtypes on host-microbe interaction. Based on this study, altered glycosylation is associated with bacterial infection. We found that adherence of *S. Typhi* is related to the abundance of host fucosylated N-glycans.

Our model provides precedence for the importance of host glycans in shaping the gut microbiota. This study will provide a great insight into the protein-carbohydrate interactions. Mass spectrometry-based approach can provide structural information on host N-glycosylation alterations and accurately identify interacting host proteins. Glycoproteomic data indicated by glycomic profile will provide and efficiently narrow down the range of potential protein targets.

FIGURES

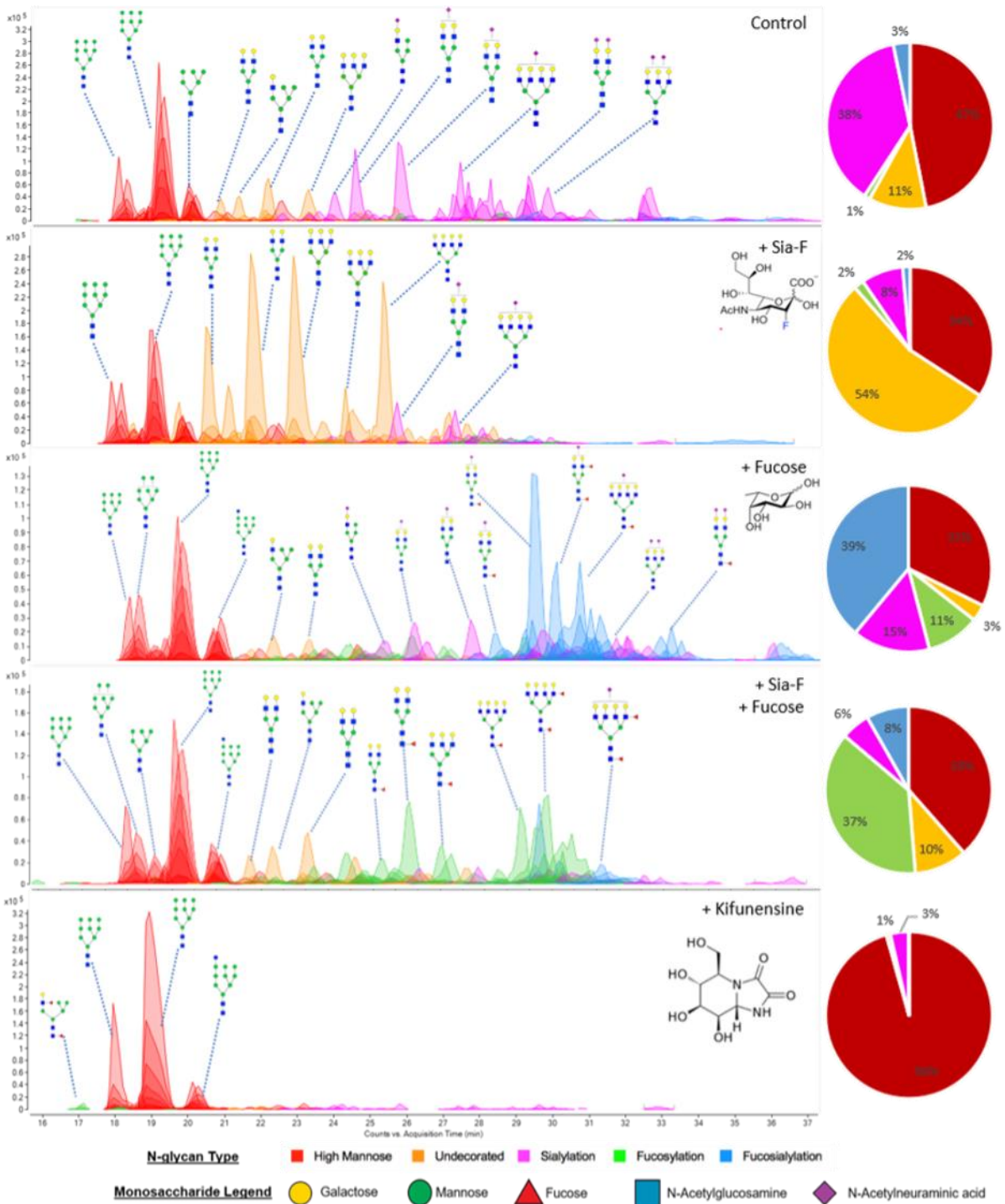


Figure 3.1. Metabolically Engineering Glycosylation in HCT116 Cell Line. Abundant peaks are annotated with putative structures. Pie charts show the summed abundance of different N-glycan types. Symbol nomenclature is used for representing glycan structures(<https://www.ncbi.nlm.nih.gov/glycans/snfg.html>)

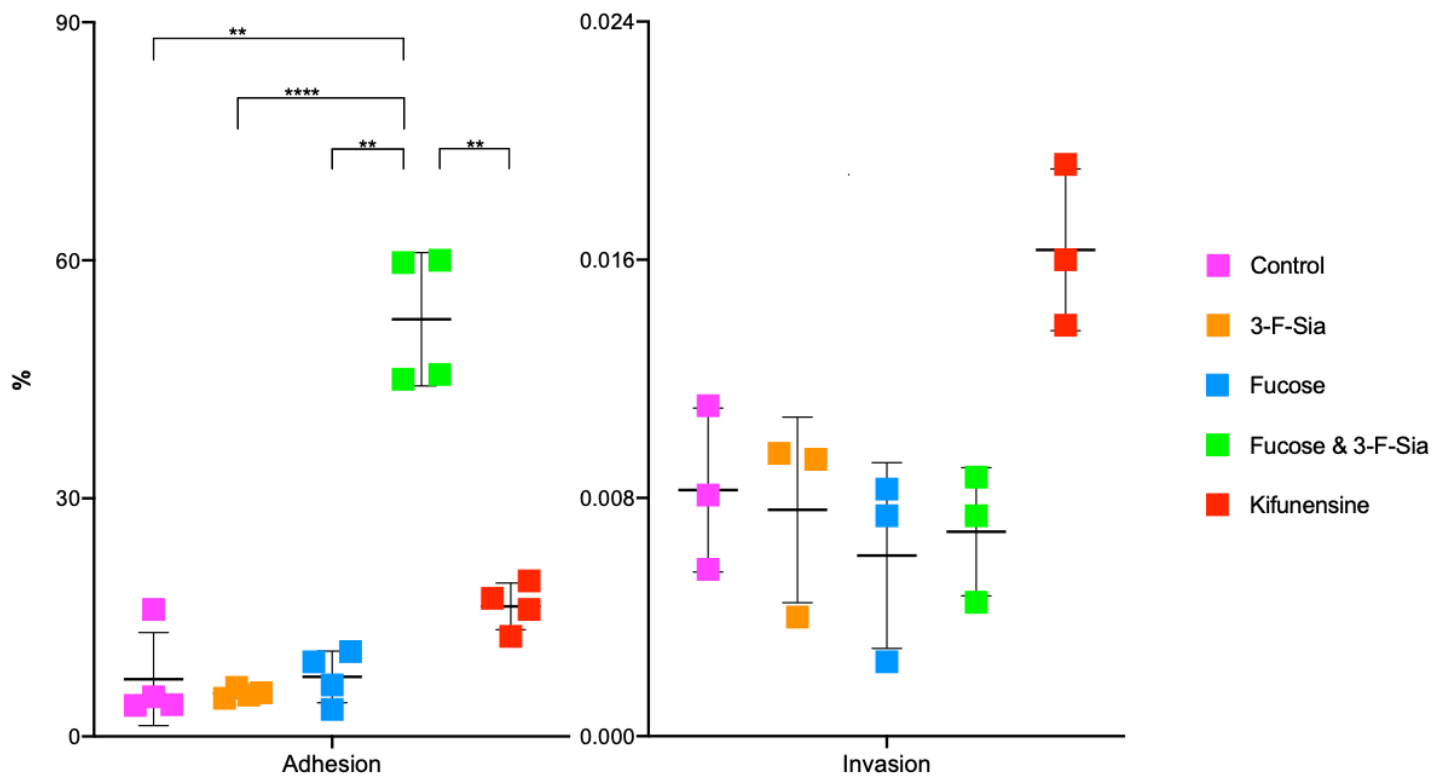


Figure 3.2. Fucosylated N-Glycans Facilitate Adherence of *Salmonella typhi* Ty2. Adherence And Invasion Assay. Data are the means and SDs from three independent experiments. Student's t-test was performed using GraphPad Prism 8 (** P < 0.01, **** P < 0.0001).

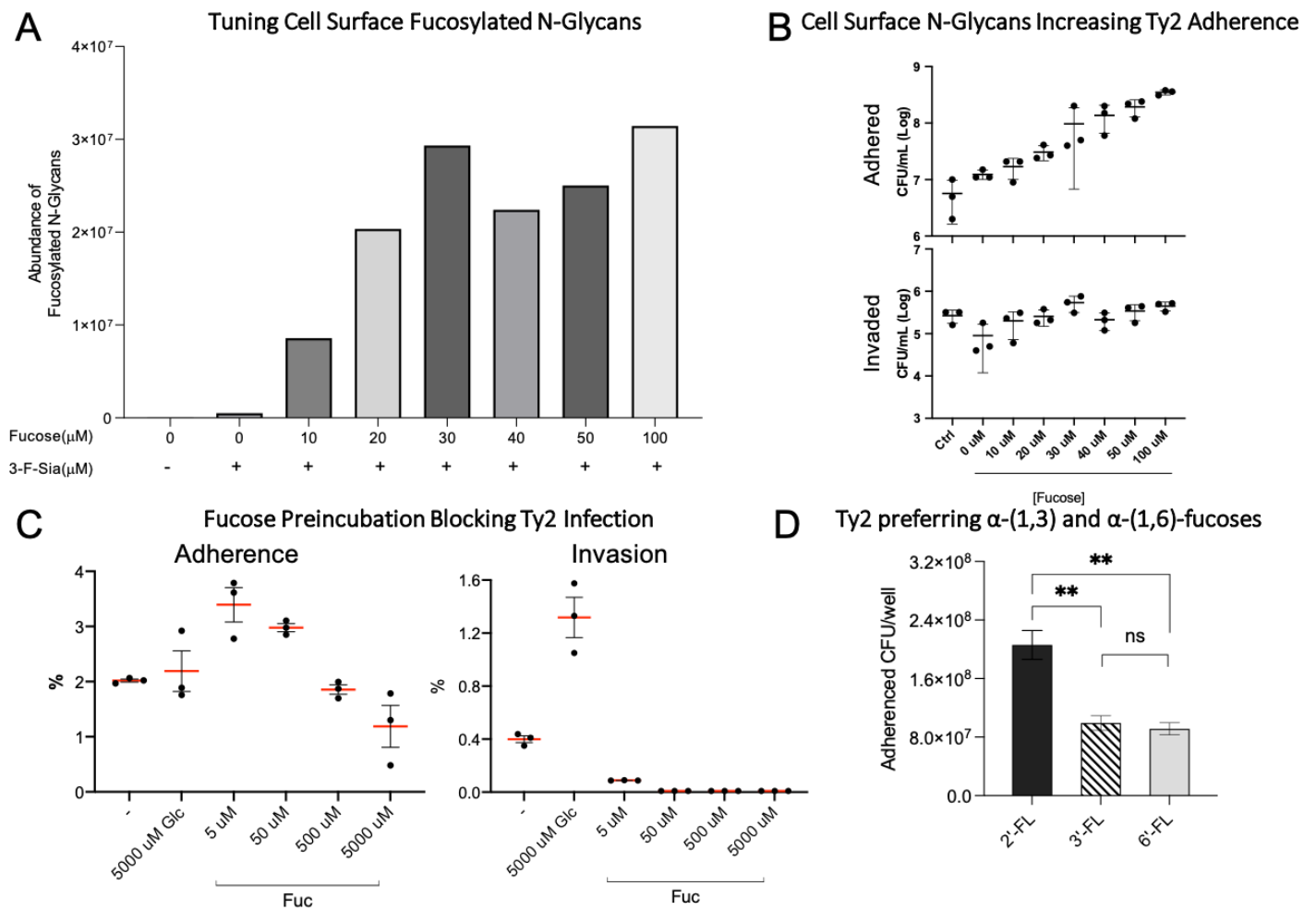


Figure 3.3. *Salmonella typhi* Ty2 Binds to Fucose Residues on Host Cell Surface. **A**. Increasing concentration of free fucose in cell media can increase the abundance of fucosylated N-glycans on the cell surface. **B, C, D**. Adherence and Invasion Assay. **B**. Adherence of *S. Typhi* to Host Cell is Associated with Fucosylated N-glycans. **C, D**. Bacteria were preincubated with carbohydrates for 60 minutes before adherence and invasion assays. Data are the means and SDs from three independent experiments. Student's t-test was performed using GraphPad Prism 8 (** P < 0.01, **** P < 0.0001).

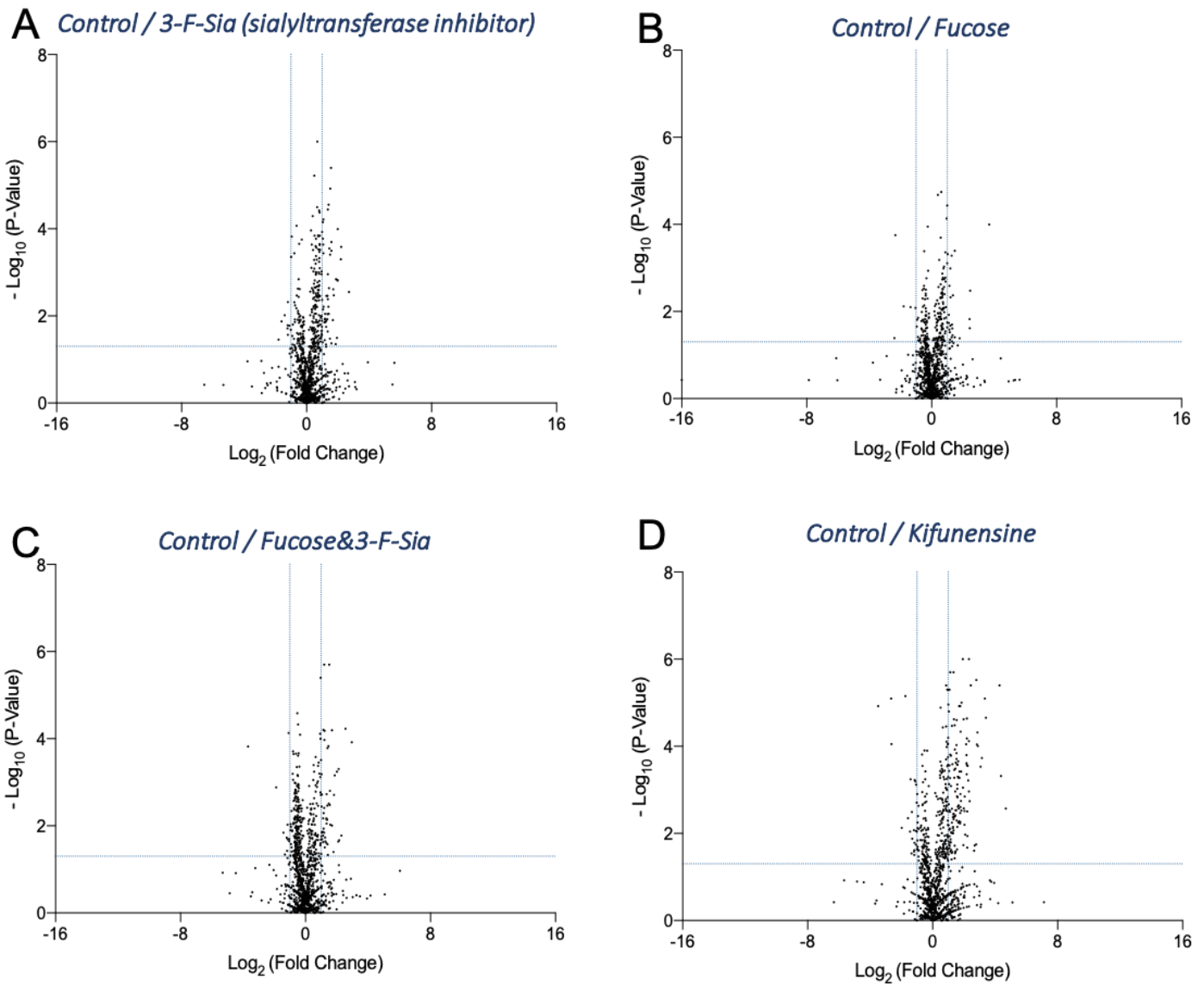


Figure 3.4. Proteomic Quantification Reveals Unchanged Expression of Most Cell Surface Proteins. **A, B, C, D** Volcano Plot of Peptide Abundance from Untreated and Treated Cells. Data are the means and SDs from three independent experiments. Significant difference: fold change >2 or <0.5, and P-Value ≤ 0.05. The student's t-test was performed using (GraphPad Software, CA).

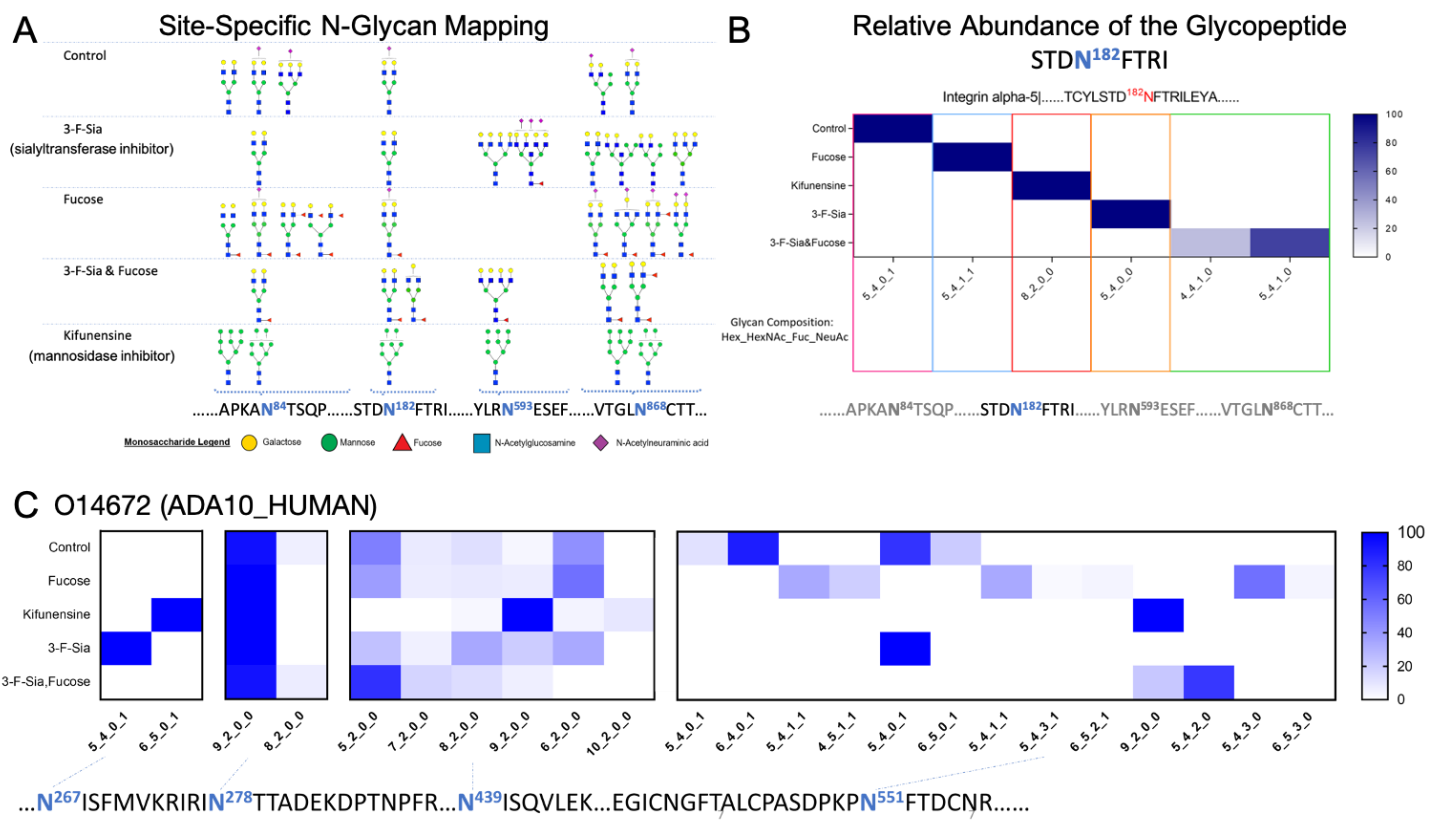


Figure 3.5. **A.** Qualitative Site Occupancy of ITA5 Protein. **B.** Relative Quantification of glycopeptide STDN¹⁸²FTRI form ITA5. **C.** Relative Quantification of glycopeptides from ADAM10.

REFERENCE

1. K. A. Karlsson, Pathogen-host protein-carbohydrate interactions as the basis of important infections. *Adv Exp Med Biol* **491**, 431-443 (2001).
2. K. Kato, A. Ishiwa, The role of carbohydrates in infection strategies of enteric pathogens. *Trop Med Health* **43**, 41-52 (2015).
3. J. Poole, C. J. Day, M. von Itzstein, J. C. Paton, M. P. Jennings, Glycointeractions in bacterial pathogenesis. *Nat Rev Microbiol* **16**, 440-452 (2018).
4. J. Holgersson, A. Gustafsson, M. E. Breimer, Characteristics of protein-carbohydrate interactions as a basis for developing novel carbohydrate-based antirejection therapies. *Immunol Cell Biol* **83**, 694-708 (2005).
5. A. Marcobal, A. M. Southwick, K. A. Earle, J. L. Sonnenburg, A refined palate: bacterial consumption of host glycans in the gut. *Glycobiology* **23**, 1038-1046 (2013).
6. J. L. Mellquist, L. Kasturi, S. L. Spitalnik, S. H. Shakin-Eshleman, The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry* **37**, 6833-6837 (1998).
7. C. R. Varki A, Esko JD, et al., editors., *Essentials of Glycobiology*, 3rd edition.
8. D. Park *et al.*, Characteristic Changes in Cell Surface Glycosylation Accompany Intestinal Epithelial Cell (IEC) Differentiation: High Mannose Structures Dominate the Cell Surface Glycome of Undifferentiated Enterocytes. *Mol Cell Proteomics* **14**, 2910-2921 (2015).
9. D. Chessa, M. G. Winter, M. Jakomin, A. J. Bäuml, Salmonella enterica serotype Typhimurium Std fimbriae bind terminal alpha(1,2)fucose residues in the cecal mucosa. *Mol Microbiol* **71**, 864-875 (2009).
10. B. Ma, J. L. Simala-Grant, D. E. Taylor, Fucosylation in prokaryotes and eukaryotes. *Glycobiology* **16**, 158R-184R (2006).
11. C. Robbe, C. Capon, B. Coddeville, J. C. Michalski, Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *Biochem J* **384**, 307-316 (2004).
12. M. Stahl *et al.*, L-fucose utilization provides Campylobacter jejuni with a competitive advantage. *Proc Natl Acad Sci U S A* **108**, 7194-7199 (2011).
13. R. Dwivedi *et al.*, L-fucose influences chemotaxis and biofilm formation in Campylobacter jejuni. *Mol Microbiol* **101**, 575-589 (2016).
14. A. Suwandi *et al.*, Std fimbriae-fucose interaction increases Salmonella-induced intestinal inflammation and prolongs colonization. *PLoS Pathog* **15**, e1007915 (2019).
15. S. Hao *et al.*, Core Fucosylation of Intestinal Epithelial Cells Protects Against. *Front Microbiol* **11**, 1097 (2020).
16. J. M. Pickard, A. V. Chervonsky, Intestinal fucose as a mediator of host-microbe symbiosis. *J Immunol* **194**, 5588-5593 (2015).
17. S. Ponniah, R. O. Endres, D. L. Hasty, S. N. Abraham, Fragmentation of Escherichia coli type 1 fimbriae exposes cryptic D-mannose-binding sites. *J Bacteriol* **173**, 4195-4202 (1991).
18. D. Chessa, C. W. Dorsey, M. Winter, A. J. Bäuml, Binding specificity of Salmonella plasmid-encoded fimbriae assessed by glycomics. *J Biol Chem* **283**, 8118-8124 (2008).
19. A. Andino, I. Hanning, Salmonella enterica: survival, colonization, and virulence differences among serovars. *ScientificWorldJournal* **2015**, 520179 (2015).
20. T. J. Griffin, A. Thanawastien, R. T. Cartee, J. J. Mekalanos, K. P. Killeen, In vitro characterization and preclinical immunogenicity of Typhax, a typhoid fever protein capsular matrix vaccine candidate. *Hum Vaccin Immunother*, (2019).
21. Q. Li, Y. Xie, G. Xu, C. B. Lebrilla, Identification of potential sialic acid binding proteins on cell membranes by proximity chemical labeling. *Chem Sci* **10**, 6199-6209 (2019).

22. Q. Li, Y. Xie, M. Wong, M. Barboza, C. B. Lebrilla, Comprehensive structural glycomic characterization of the glycocalyxes of cells and tissues. *Nat Protoc* **15**, 2668-2704 (2020).
23. K. Nakayama *et al.*, Mutation of GDP-mannose-4,6-dehydratase in colorectal cancer metastasis. *PLoS One* **8**, e70298 (2013).
24. N. Mishra, M. Spearman, L. Donald, H. Perreault, M. Butler, Comparison of two glycoengineering strategies to control the fucosylation of a monoclonal antibody. *J Biotechnol* **324S**, 100015 (2020).
25. S. Li *et al.*, Glycoengineering of Therapeutic Antibodies with Small Molecule Inhibitors. *Antibodies (Basel)* **10**, (2021).
26. Y. Narimatsu *et al.*, Genetic glycoengineering in mammalian cells. *J Biol Chem* **296**, 100448 (2021).
27. K. Macharoen *et al.*, Effects of Kifunensine on Production and. *Int J Mol Sci* **21**, (2020).
28. D. Park *et al.*, Salmonella Typhimurium Enzymatically Landscapes the Host Intestinal Epithelial Cell (IEC) Surface Glycome to Increase Invasion. *Mol Cell Proteomics* **15**, 3653-3664 (2016).
29. A. Bren *et al.*, Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP. *Sci Rep* **6**, 24834 (2016).
30. A. Porter *et al.*, A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology* **20**, 369-380 (2010).
31. J. Heimbürg-Molinaro, X. Song, D. F. Smith, R. D. Cummings, Preparation and analysis of glycan microarrays. *Curr Protoc Protein Sci* **Chapter 12**, Unit12.10 (2011).
32. C. D. Rillahan, J. C. Paulson, Glycan microarrays for decoding the glycome. *Annu Rev Biochem* **80**, 797-823 (2011).
33. Y. Kobayashi, H. Tateno, H. Ogawa, K. Yamamoto, J. Hirabayashi, Comprehensive list of lectins: origins, natures, and carbohydrate specificities. *Methods Mol Biol* **1200**, 555-577 (2014).
34. M. Wührer, M. I. Catalina, A. M. Deelder, C. H. Hokke, Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J Chromatogr B Analyt Technol Biomed Life Sci* **849**, 115-128 (2007).
35. M. Wührer, C. A. Koeleman, C. H. Hokke, A. M. Deelder, Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Commun Mass Spectrom* **20**, 1747-1754 (2006).

SUPPLEMENTARY INFORMATION

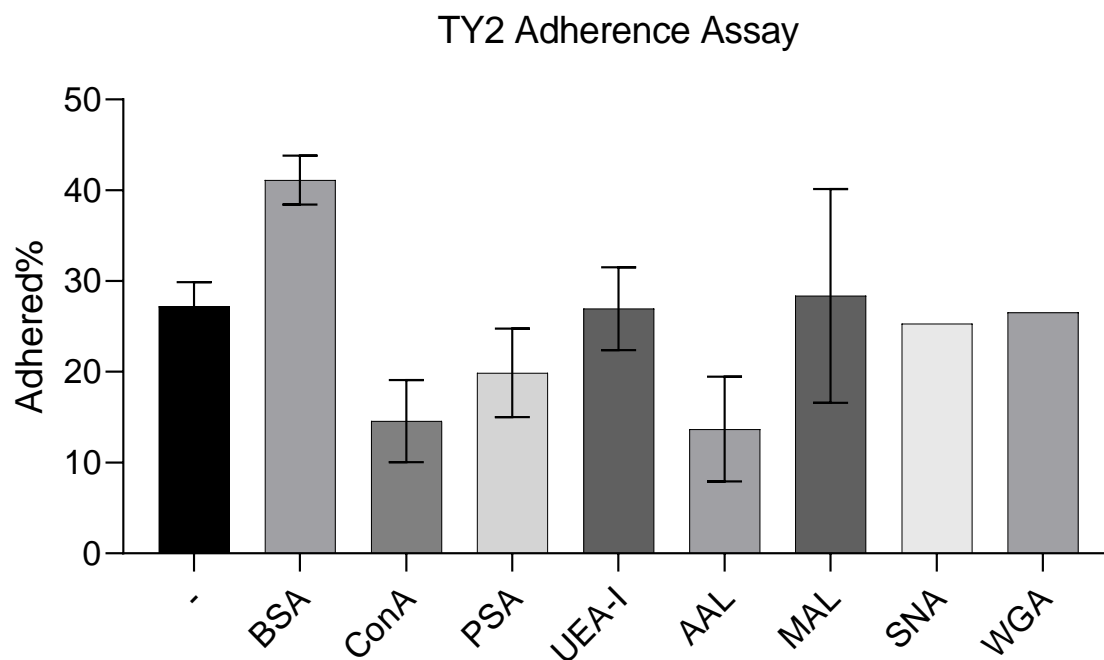


Figure S 3.1. Lectin Preincubation Confirms the Preference of Ty2 for α -(1,3) and α -(1,6)-fucoses.

Table S 3.1. Summary of Lectin Specificity

Lectin	Reported glycan selectivity
ConA	High-Mannose including Man α 1-6(Man α 1-3)Man
PSA	Fuc α 1-6GlcNAc (Core Fuc) , α -Man
UEA-I	Fuc α 1-2Gal β 1-4GlcNAc (H-type 2)
AAL	Fuc α 1-3(Gal β 1-4)GlcNAc (Lewis x), Fuc α 1-6GlcNAc (Core Fuc)
MAL_I	Sia α 2-3Gal β 1-4GlcNAc
SNA	Sia α 2-6Gal/GalNAc
WGA	(GlcNAc β 1-4) _n (Chitin), Hybrid type N-glycan, Sia

Chapter 4 Determination of the glycoprotein specificity of lectins on cell membrane through oxidative proteomics

Authors:

Yixuan Xie¹; Ying Sheng²; Qiongyu Li¹; Seunghye Ju¹; Joe Reyes⁴; Carlito B Lebrilla^{1,3}

¹Department of Chemistry, University of California, Davis, Davis, California, United States.

²Department of Chemistry, Biochemistry, Molecular, Cellular and Developmental Biology Graduate Group, University of California, Davis, Davis, California, United States. ³Department of Biochemistry, University of California, Davis, Davis, California, United States. ⁴Marine Science Institute, University of the Philippines, Diliman, Quezon City, Philippines. Correspondence and requests for materials should be addressed to C.B.L. (email: cblebrilla@ucdavis.edu)

ABSTRACT

The cell membrane is composed of a network of glycoconjugates including glycoproteins and glycolipids that presents a dense matrix of carbohydrate playing critical roles in many biological processes. Lectin-based technology have been widely used to characterize glycoconjugates in tissues and cell lines. However, their specificity toward their putative glycan ligand and sensitivity *in situ* has been technologically difficult to study. Additionally, because they recognize primarily glycans, the underlying glycoprotein targets are generally not known. In this study, we employed lectin proximity oxidative labeling (Lectin PROXL) to identify cell surface glycoproteins that contain glycans that are recognized by lectins. Commonly-used lectins were modified with a probe to produce hydroxide radicals in the proximity of the labeled lectins. The

underlying polypeptide of the glycoproteins recognized by the lectins are oxidized and identified by the standard proteomic workflow. As a result, approximately 70% of identified glycoproteins were oxidized *in situ* by all the lectin probes, while only 5% of the total proteins were oxidized. The correlation between the glycosites and oxidation sites demonstrated the effectiveness of lectin probes. The specificity and sensitivity of each lectin were determined using site-specific glycan information obtained through glycomic and glycoproteomic analyses. Notably, the sialic acid-binding lectins and the fucose-binding lectins have higher specificity and sensitivity compared to other lectins, while those that were specific to high mannose glycans has poor sensitivity and specificity. This method offers unprecedented view of the interactions of lectins with specific glycoproteins as well as protein networks that are mediated by specific glycan types on cell membranes.

INTRODUCTION

The carbohydrate layer on the cell surface that is anchored by protein and lipid scaffolds is involved in a host of important and central cellular processes, including cellular adhesion, immune defense, and cell permeability.(1-3) Extensive covalent and non-covalent interactions occur on the cell membrane that defines the topology and availability of glycan-binding sites and antigens. Cell-binding proteins such as cadherins and integrins are glycosylated with glycans centrally involved in their function.(4, 5) Enzymes used for protection such as DPP IV (dipeptidyl peptidase IV) and IAP (alkaline phosphatase) are also glycosylated with their function affected by glycosylation.(6)

Lectin-based techniques, including lectin microarray, lectin-affinity enrichment, and enzyme-linked lectin assay (ELLA) are extensively used for studying specific glycan structures *in vitro*.(7) However, due to the transient and weak nature of glycan-mediated interactions, glycans need to be in their native environment to resolve the structure and dynamics of such interaction. Various glycoprotein models have also been introduced to mimic the spatial distribution of glycan epitopes on natural glycoconjugate ligands.(8-10) Electrochemical imaging have been widely employed to visualize and characterize glycoconjugates *in situ*.(11) Recently, Han *et al.* developed a method involving using laser cleavable lectin probes for glycan detection at single cell level through mass spectrometry (MS)-based analysis.(12) These methods rely highly on the putative specificities of lectins towards glycoconjugates, however the in-situ targets of the lectins are generally unknown thereby hindering the broader applications of these methods.

Several strategies have been developed to correlate the visual information with detailed protein information. In particular, metabolic labeling of bioorthogonal reporters has been

introduced to study the in-situ interactions between lectins and glycans. Paulson and co-workers applied the photocrosslinking sialic acid to identify the *cis*- and *trans*- targets of Siglec-2 *in situ*.(13, 14) Kohler and co-workers complemented the method with photocrosslinking sugars modified at C-5 position.(15) However, these investigations were focused on specific lectin-glycoprotein partners. Oxidative labeling methods have been used to examine protein-protein pairwise interactions. The oxidative reagent, iron (S)-1-(p-bromoacetamidobenzyl) EDTA (FeBABE), was created and used to identify protein-associated interactions.(16) Oxidative labeling has been used more broadly to obtain interactive relationships, as with chemical cross-linking, but in a more generalized manner by providing proximity information of proteins involved.(17)

In this research, we used oxidative labeling by reacting a lectin with a dibenzocyclooctyne-FeBABE (DBCO-FeBABE) as a probe to generate hydroxyl radicals (**Figure 1**). The primary amine on the lectin was functionalized with azide, followed by conjugation to the azido group to form oxidative probes on the modified lectins (**Figure 2a**). The probe produced hydroxyl radicals when hydrogen peroxide was introduced. Thus, proteins near the lectins were oxidized by the generated hydroxyl radicals and the oxidatively modified side chains were characterized using nanoLC-MS. The lectin proximity oxidative labeling (Lectin PROXL) method extends a previous one mapping potential sialic acid-associating protein on the cell surface.(18) Eight commonly used lectins were chosen to identify the specific glycoprotein targets of each lectin (**Table 1**). *Sambucus nigra* agglutinin (SNA) binds to sialic acid with $\alpha(2,6)$ linkage preference, and *Maackia amurensis* leucoagglutinin (MAL) binds to $\alpha(2,3)$ sialic acid.(19, 20) *Aleuria aurantia* lectin (AAL) binds to fucose in general, while *Pisum sativum* agglutinin (PSA) prefers core fucosylation.(21)

Both *Phaseolus vulgaris* leucoagglutinin (PHA-L) and *Phaseolus vulgaris* erythroagglutinin (PHA-E) bind to galactose, but PHA-L prefers tri/tetra-antennary N-glycans and PHA-E have a higher affinity toward bi-antennary N-glycan.(22, 23) Hipppeastrum hybrid lectin (HHL) binds to high mannose type N-glycans through recognizing $\alpha(1,3)$ and $\alpha(1,6)$ mannose.(24) Wheat germ agglutinin (WGA) with N-acetylglucosamine and sialic acid binding properties was chosen to investigate the general targets.(25) Lectin PROXL was applied to evaluate the interactions between lectins and glycoproteins to provide the protein targets on the cell surface, and the functional analysis of specifically oxidized proteins provided networks that were mediated by glycans on cell membrane.

METHODS AND MATERIALS

Modification of lectin with Fe(III) probe

The lectin was dissolved in PBS at a final concentration of 1mg/mL. 10 μ L of NHS-PEG4-Azide (100 mM) was added to the lectin solution. The reaction carried at room temperature for 1 hour, followed by adding 10 μ L of Tris buffer (1M) to quench the reaction and the excess reagent was removed using the ultra-centrifugal filter. DBCO-FeBABE was prepared in advance, and detailed procedures for synthesizing DBCO-FeBABE was described previously.⁽¹⁸⁾ 10 μ L of DBCO-FeBABE (100 mM) was added to the mixture, and the reaction was carried in a water bath at 37 °C for 4 hours. The oxidation probe-modified lectin was further purified using the 10k ultra-centrifugal filter.

Proteins and reagents

PNT2 and LNCaP cell lines were obtained from the American Type Culture Collection (ATCC). Lectins were purchased from Vector Laboratories. FeBABE was purchased from Dojindo Molecular Technologies. Dithiothreitol (DTT), iodoacetamide (IAA), DBCO-NH₂, and DBCO-Cy3 were purchased from Sigma-Aldrich. Phosphate Buffered Saline (PBS), Roswell Park Memorial Institute (RPMI) 1640 medium, fetal bovine serum (FBS), penicillin, NHS-PEG4-Azide, and Hoechst 33342 were purchased from ThermoFisher Scientific. Sequencing Grade Modified Trypsin was purchased from Promega.

Cell culture

Human immortalized normal prostatic epithelial PNT2 cells and human prostate carcinoma epithelial cells LNCaP were obtained from ATCC and grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% (v/v) fetal bovine serum (FBS), 1% (v/v)

penicillin. Cells were maintained in a humidified incubator at 37 °C with 5% CO₂ and subcultured at 80% confluency.

Confocal imaging of lectin

Human immortalized normal prostate epithelial PNT2 cells were obtained from ATCC and cultured in FluoroDish™ cell culture dishes (WPI, FL). Around 60% confluency, cells were fixed with 4 % paraformaldehyde. 5 µg of fluorescein-conjugated lectin or DBCO-Cy3-conjugated lectin was added to the cells, and the labeling of cell surface glycans was performed at 37 °C for 30 minutes, followed by staining the nucleus with Hoechst 33342 at 37 °C for 10 minutes. Confocal images were captured using a Leica TCS SP8 STED 3X Super-Resolution Confocal Microscope (Wetzlar, Germany). The images were analyzed and processed using Imaris software (Oxford Instruments, Switzerland)

Oxidation of cell surface glycoprotein

For oxidative mapping of the sialic-acid environment on the cell surface, 3 x 10⁶ cells were treated with serum-free media supplemented with 20 µg of the lectins modified with the oxidation probe for 30 minutes at 37 °C. The cells were washed with PBS three times to wash off unbound lectins, followed by the treatment with 100 µM hydrogen peroxide for 30 minutes at 37 °C. The hydroxyl radicals were quenched with 10 mL of 10 mM methionine amide hydrochloride in PBS. Cells were harvested and resuspended in a homogenization buffer containing protease inhibitor cocktail (EMD Millipore, CA), 0.25 M sucrose, and 20 mM HEPES–KOH (pH 7.4). Cells were lysed at 4 °C using a probe sonicator (Qsonica, CT) performed with alternating on and off pulses of 5 and 10 seconds, respectively.

Cell membrane extraction

Detailed procedures for cell membrane extraction as described previously.⁽³⁷⁾ Briefly, nucleus and cellular debris were isolated by centrifugation at 2,000 x g for 10 minutes. Mitochondrial fraction was removed by centrifugation at 12,000 x g for 10 minutes. The supernatants were subjected to ultra-centrifugation at 200,000 x g for 45 minutes to extract the membrane fraction. The crude membrane fraction was further washed with 500 μ L of Na₂CO₃ (0.2 M) and nanopure water, respectively.

Proteins digestion and purification

The cell membrane pellets were reconstituted with 60 μ L of 8 M urea and sonicated for 15 minutes for denaturing. 2 μ L of DTT (550 mM) was added to the samples and incubated for 50 minutes at 55 °C, and the free cystine was alkylated with 4 μ L of IAA (450 mM) 20 minutes at room temperature in the dark. 30 μ L of trypsin (0.1 μ g/ μ L) was added to the mixture and the pH of the solution was adjusted using 420 μ L of ammonium bicarbonate (3.95 mg/mL). The tryptic digestion was performed at 37 °C for 18 hours. The resulting peptides were purified using solid-phase extraction with C18 cartridges. The tryptic digestion of glycoproteomics samples was prepared the same as proteomic analysis, iSPE[®]-HILIC cartridges (The Nest Group, MA) was used to enrich the glycopeptides.

Proteomic analysis using LC-MS/MS

The proteomics and glycoproteomics samples were characterized using UltiMate[™] WPS-3000RS nanoLC system coupled with Orbitrap Fusion Lumos (ThermoFisher Scientific). 1 μ L of the sample was injected, and the analytes were separated on Acclaim[™] PepMap[™] 100 C18 LC Column (3 μ m, 0.075 mm x 250 mm, ThermoFisher Scientific) at a flow rate of 300 nL/min. Water

containing 0.1% formic acid and 80% acetonitrile containing 0.1% formic acid were used as solvents A and B, respectively. MS spectra were collected with a mass range of m/z 600–2000 at a rate of 1.5 s per spectrum in positive ionization mode. The filtered precursor ions in each MS spectrum were subjected to fragmentation through 30% higher-energy C-trap dissociation (HCD) with nitrogen gas.

Data Analysis

Oxidized proteins and glycoproteins were identified using Byonic software (Protein Metrics, CA) against the human protein database (UniProt). Alkylation of cysteine with carbamidomethylation was assigned as a fixed modification. Deamidation of asparagine and glutamine, methylation of lysine and arginine, and acetylation of the protein N-terminus were assigned as the rare variable modification. For oxidized samples, the oxidized modification was selected as common variable modifications according to previous settings. For glycoprotein samples, an in-house human N-glycan database was applied for asparagine N-glycosylation. The unmodified and modified peptides were then quantified using Byologic. The extent of modification was calculated by dividing the abundance of the modified peptide to the total abundance of corresponding wildtype and modified peptide.

Cell surface N-glycomic analysis

The cell membrane fractions were resuspended with 100 μ L of 5 mM DTT in 100 mM ammonium bicarbonate. The mixture was heated in boiling water for 3 minutes. The cleavage of N-glycans was performed by adding 2 μ L of PNGase F followed by the incubation at 37 °C water bath overnight. The released N-glycans were separated using 200,000 \times g for 30 minutes, and the supernatant was purified using porous graphitic carbon (PGC) on an SPE plate. The glycan

samples were dried and reconstituted in 30 μ L nanopure water. 5 μ L of the sample was injected and analyzed with an Agilent 6520 Accurate Mass Q-TOF LC/MS equipped with a PGC nano-chip (Agilent, CA). A binary gradient using solvent A with 3% (v/v) ACN and 0.1% (v/v) formic acid in water and solvent B with 90% (v/v) ACN and 1% (v/v) formic acid in water was applied to separate N-glycans at 300 nL/min flow rate. The resulted chromatographs of glycans were extracted with the MassHunter Qualitative Analysis B08 software (Agilent, CA). N-glycan compounds were identified with an in-house library that contains the accurate mass and formula of human N-glycans, and the N-glycan structures were confirmed through tandem MS fragmentation.

RESULTS

Production of Lectin PROXL Probes

To confirm the efficacy of overall the reaction, bovine serum albumin (BSA) was used as a model protein and characterized by nanoLC-MS. The modification of the protein with azide-PEG4-NHS was performed and the protein subjected to conjugation of the probe. As shown in **Figure S1**, the tandem MS/MS data confirmed the presence of the azido group modification on BSA (+273.13 Da), as well as the conjugation of DBCO (+549.26 Da) on K437 residue. A greater than 80% conversion was achieved for this site as determined by the intensities of modified over wild-type peptides. Additionally, there were two other sites that reacted and yielded an incorporation of 60 to 70%. Missed cleavages were observed at the modified lysine residues.

To determine whether the binding properties of lectins were affected by the modifications, we prepared DBCO-cy3-labeled SNA using DBCO-cy3 and applied it towards PNT2 cells. SNA has specificities towards $\alpha(2,6)$ sialic acids. Cells were fixed with formaldehyde, treated with the lectin, and then stained with Hoechst to observe the nucleus. No significant decrease in the fluorescent intensity was observed in the modified lectin thereby indicating that the modification did not alter the interactions between the lectin and its receptors on the cell surface (**Figure S2a and b**).

Oxidative Labeling of Glycoproteins on Cell Surfaces of PNT2 Cell Lines

The condition for oxidative labeling was optimized by varying the hydrogen peroxide concentrations and the reaction times, which governs the flux and diffusion distances of the hydroxyl radicals. Higher concentrations of hydrogen peroxide yielded longer distances for the radicals to travel, while low concentrations yielded shorter distances. Previous studies on the

effects of concentrations showed the radicals can travel as far as 60 Å.(26) We also previously found that 50-300 μM hydrogen peroxide concentrations and 30 minutes reaction time produce hydroxyl radicals within 20 Å of the probe, which was optimal for identifying proteins that were in the vicinity of the probe.(18) In these experiments, we also found 30 minutes reaction time to be optimal and varied the hydrogen peroxide concentrations from 50 μM to 300 μM. We used WGA for optimization because WGA binds putatively to nearly all glycans through its specificity for N-acetylglucosamines (GlcNAc) and sialic acids.

The results indicated that the number of oxidized proteins increased with the higher hydrogen peroxide concentrations, but did not increase significantly more at concentrations greater than 100 μM. To monitor the extent of the reactions, we determined the increase in the number of oxidation sites as well as the extent of oxidations on specific sites (**Figure S3**). The subcellular locations of oxidized proteins were annotated using STRING software and most of the oxidized proteins were localized at the cell plasma membrane. Background oxidation was identified and noted through the use of control experiments. **Figure S4** shows that sites that were oxidized in the control and further increased with increasing hydrogen peroxide concentrations. These proteins were nonglycosylated and were readily eliminated from the analysis. Based on these results, we chose 100 μM and 30 minutes as the optimal reaction conditions for Lectin PROXL. The results also confirmed that nonspecific oxidation can be monitored and overoxidation avoided by limiting the reaction time and hydrogen peroxide concentrations.

To monitor the overall oxidation of the proteins, standard proteomic analysis was performed on cells. For PNT2 cells, the experiments typically yielded more than 200 oxidized proteins for all eight lectins examined (**Supporting Data 1-2**). Approximately 70% of identified

glycoproteins were oxidized by the lectin probe, while only 5% of the nonglycosylated proteins were oxidized. Other general observations were made. Methionine was the most commonly oxidized amino acid. Other amino acids including cysteine, tryptophan, and phenylalanine were oxidized but to a significantly lower degree. The proteins that were oxidized consisted of both glycosylated and nonglycosylated proteins (**Figure 2b**). The glycosylated proteins contained the target glycans as shown below. The oxidized nonglycosylated proteins were those that interacted with the target glycoproteins.

We further observed a general correlation between the number glycosylation sites and oxidation of glycoproteins. Of the glycoproteins containing only one glycosylation site, over 35% were oxidized. With those containing more than one glycosite, over 75% of these glycoproteins were oxidized. Notable examples include CD166 (CD166 antigen) with N-glycans at N167, N265, N361, N480 was oxidized by SNA probe. Conversely, the protein TSN13 (tetraspanin-13) with sialofucosylated glycans at a single site N137 was not oxidized by any of the lectins.

Determination of the Relationship Between Sites of Glycosylation and Sites of Oxidation

To obtain the site-specific information of oxidized glycoproteins, including peptide sequences, glycosites, and glycan compositions, we employed a secondary glycoproteomic analysis of the cell membrane proteins (shown in **Supporting Data 3**). The membrane proteins were digested, and the glycopeptides enriched through hydrophilic interaction chromatography (HILIC) using solid-phase extraction (SPE). The enriched fractions were analyzed using nanoLC-MS and the sites of oxidation were examined relative to the sites of glycosylation to determine the spatial relationship between them. For example, SNA is a lectin known to recognize sialylated glycans. Sialylated glycans including $\text{Man}_{(3)}\text{Gal}_{(2)}\text{GlcNAc}_{(4)}\text{Sia}_{(1)}$, $\text{Man}_{(3)}\text{Gal}_{(2)}\text{GlcNAc}_{(4)}\text{Sia}_{(2)}$, and

Man₍₃₎Gal₍₃₎GlcNAc₍₅₎Sia₍₃₎, were found specifically at N343 on ITA2 (integrin alpha-2), while high-mannose type glycans Man₍₈₎GlcNAc₍₂₎ and Man₍₇₎GlcNAc₍₂₎ were found at N432 (**Figure S5**). From the oxidation results, we found that ITA2 was oxidized by SNA lectin at T337, which was closer to the site of sialylation than mannosylation.

AMPN (aminopeptidase N) is a highly glycosylated protein with N-glycans distributed over four glycosites including N128, N234, N265, and N681 (**Figure 3a**). The SNA probe oxidized the protein extensively at M354, M435, M444, V632, and M693, which were all near the glycosites associated with sialylated glycans. Another sialic acid recognizing lectin, MAL with the putative specificity for $\alpha(2,3)$ sialic acid labeled the protein only at M693. These results suggested that the sialic acids on N265 are likely $\alpha(2,6)$ sialic acid due to the proximity, while the sialylated glycans at N681 likely contained both $\alpha(2,6)$ and $\alpha(2,3)$ sialic acids due to the oxidation at M693 for both probes. High mannose N-glycans were also detected at N128 exclusively. Indeed, the high mannose recognizing lectin HHL yielded oxidation of AMPN at F103. Glycoproteomic analysis also revealed fucosylated glycans at the same four glycosites. Similarly, the oxidation results were consistent with the localization by the fucose-binding lectin AAL at Y161, M199, M444, M486, and M693. Interestingly, the galactose-binding lectin, PHA-E, oxidized the AMPN protein only at M486 and M693, although bi-antennary glycans were observed at all four glycosites. The WGA probe was expected to oxidize the glycoprotein at the same sites as all the other lectins. WGA is generally believed to bind to all N-glycans through its affinity with the GlcNAc and sialic acids. As a result, the WGA lectin probe was found to oxidize AMPN protein at F103, Y161 and M444, M496, and M693 consistent with expectations.

Not all lectins were equally effective with binding to their putative targets. The lectin PSA is specific towards core fucosylation, however despite the large number of fucosylated glycans AMPN was not oxidized by PSA. The lack of reactivity could be attributed to two factors. One is that these glycans contained no core fucosylation. Or, the core fucose was deep within the fold of the protein and could not be accessed by the lectin. As core fucosylation is much more common than antenna fucosylation, we believe the latter is more correct. Below, we show modeling calculations that exhibit this behavior. Another lectin, PHA-L, with a specificity towards galactose on termini of tri- and tetra-antennary did not yield oxidized AMPN products, despite the presence of tri- and tetra-antennary structures. In these glycans, the termini contain fucose and sialic acids, and the presence of these residues likely block the terminal galactose from binding with the lectin, the galactose. It has been shown that the affinity of this lectin toward galactose is diminished by the presence of the sialic acid and fucose on the galactose.(27)

Glycoprotein specificity of lectins

To determine the specificity of the lectins, we determined the number of glycoproteins oxidized and the fraction that contains the putative lectin target (**Figure 3B**). For example, all 48 glycoproteins that were oxidized by SNA contained sialylated glycans (100% specificity), while only 35% (7 out of 20) of glycoproteins oxidized by HHL contained high-mannose type glycans. For clarity, we grouped the lectins into four types, namely sialylated-binding lectins (SNA and MAL), fucosylated-binding lectins (AAL and PSA), undecorated glycan-binding lectins (PHA-E, PHA-L, and HHL) and a glycan-binding lectin with broad specificities (WGA).

The lectins SNA and MAL both bind glycans with sialic acids, however SNA is specific for $\alpha(2,6)$ while MAL for $\alpha(2,3)$. From the proteomic analysis, all glycoproteins oxidized by SNA were

sialylated as were all 20 by MAL. The larger number of glycoproteins marked by SNA compared to MAL implied the greater presence of $\alpha(2,6)$ versus $\alpha(2,3)$ in PNT2 cells. These results were further validated for this cell line with fluorescence labeling, which showed greater fluorescence with SNA compare to MAL (**Figure S2c**). Of the glycoproteins oxidized by MAL, 17 were also oxidized by SNA suggesting the presence of both linkages in those proteins (**Figure S6a**). The three glycoproteins uniquely oxidized by MAL suggested that these sialylated glycoproteins contained primarily $\alpha(2,3)$ sialic acid. To confirm this notion, we treated the glycoprotein from PNT2 cells with $\alpha(2,3)$ sialidase. As shown in **Supporting Data 4**, one of the proteins, MPRD (protein cation-dependent mannose-6-phosphate receptor) was found to have a sialylated glycan at N83 corresponding to Hex₍₅₎HexNAc₍₄₎Fuc₍₁₎NeuAc₍₁₎. Treatment of the glycopeptide mixture with the sialidase resulted in the loss of the sialic acid and the appearance of the desialylated glycopeptide confirming the linkage of this glycoform. Other glycopeptides belonging to the three proteins yielded the similar results.

The fucose-binding lectins, AAL and PSA, yielded 45 and 23 oxidized proteins, respectively. AAL has broad specificities towards core and antenna fucose, while PSA prefers mainly to core fucose. 21 glycoproteins marked by PSA were found in AAL proteins (**Figure S6b**). The oxidized glycoproteins observed in common were found to have monofucosylated glycans, while the glycoproteins that were uniquely labeled by AAL were dominated by difucosylated and trifucosylated structures that likely contain fucose at the antenna. Not all glycoproteins with the monofucosylated glycans were labeled by PSA. For example, protein ITGB1 (integrin beta-1) with Hex₍₅₎HexNAc₍₄₎Fuc₍₁₎Sia₍₂₎ at N97 was oxidized at L108, while EGFR (epidermal growth factor receptor) with the same glycan composition at N528 did not yield oxidized peptides. However,

EGFR contained multiplyfucosylated and were oxidized by AAL. There are at least two reasons for why EGFR was not oxidized by PSA: the lone fucose was not at the core, or the core fucose was not accessible due to steric shielding. To investigate the accessibility of the core fucose on these two proteins, we built two glycoprotein models and minimized the energy using Glycam.(28) As shown in **Figure 4**, fucose on EGFR was predicted to be sterically obscured by K538 and L541 side chains in the vicinity of N-glycan site. Conversely, the core fucose on ITGB1 was unhindered and readily accessible. The results suggest that protein folding and spatial accessibility of glycans could be a factor that hinder recognition by lectins.

Galactosylated structures are putatively recognized by lectins PHA-L and PHA-E, albeit the two lectins have different affinities towards various number of antennas. PHA-L favors tri/tetra-antennary, while PHA-E binds with bi-antennary structures. We found glycoproteins uniquely oxidized by PHA-E indeed contained mainly bi-antennary N-glycans, while tri- and tetra-antennary N-glycans were the primary targets of the PHA-L probe. In general, PHA-E yielded more oxidized glycoproteins, which suggested the cell line had more bi-antennary instead of tri/tetra-antennary structures. We further employed N-glycomic analysis on PNT2, and we found that the cells indeed had more bi-antennary N-glycans such as $\text{Man}_{(3)}\text{Gal}_{(2)}\text{GlcNAc}_{(4)}\text{Fuc}_{(1)}$ and $\text{Man}_{(3)}\text{Gal}_{(2)}\text{GlcNAc}_{(4)}\text{Fuc}_{(1)}\text{Sia}_{(1)}$ than higher antennary structures (**Figure S7**).

Nearly all the lectins had very high specificity towards the putative protein-associated glycan with perhaps one exception, namely HHL. HHL is commonly used to identify high mannose type N-glycans, because it can recognize both (α 1,3) and (α 1,6) mannose structures. Due to the lack of high-mannose type glycans on the PNT2 cell surface, HHL only generated 20 oxidized glycoproteins. Within this group, only seven glycoproteins were found to have high-mannose

structures. This apparent lack of specificity suggests that HHL binding is perhaps not limited to high-mannose type glycans, it has been shown that HHL can also bind to N-glycans with terminal galactose and sialic acid.(29) Therefore, caution should be taken when HHL is used strictly for monitoring the amount of high mannose structures.

WGA is a lectin with broad specificities and is widely used as a probe for monitoring all N-glycans. Probing PNT2 with WGA yielded over 50 oxidized glycoproteins. Indeed, WGA yielded the highest number of oxidized glycoproteins among all eight lectins investigated. It should be further noted that the glycoproteins oxidized by WGA probe were also oxidized with other lectin probes thereby confirming the broad specificity of WGA towards N-glycans. However, not all glycoproteins oxidized by the other lectins were oxidized by the WGA probe. To determine how WGA differentiates glycoproteins, we employed glycoproteomic analysis of the cell line. We found that the glycoproteins not oxidized by WGA contained primarily complex type N-glycans. For example, the complex type N-glycan, Hex₍₅₎HexNAc₍₄₎Fuc₍₂₎NeuAc₍₁₎, was found on N166 of ECE1 (endothelin-converting enzyme 1). While the parent protein was oxidized by AAL, it was not marked by WGA. On the other hand, most of WGA oxidized glycoproteins contained hybrid-type structures. For example, the protein EPCAM (epithelial cell adhesion molecule), which was oxidized by WGA, contained hybrid N-glycans, Man₍₄₎Gal₍₁₎GlcNAc₍₃₎Sia₍₁₎ and Man₍₄₎Gal₍₁₎GlcNAc₍₃₎Fuc₍₁₎, on N111. Although WGA is commonly used to monitor generally N-glycosylation, our results suggested that WGA prefers the hybrid type of N-glycans over the complex type N-glycans.

We further determined the sensitivity of Lectin PROXL by examining the putative target glycoproteins and comparing them to those that were subsequently oxidized (**Figure 3c**). The

fraction of glycoproteins oxidized by the probe revealed the general sensitivity of the lectin, which was generally determined to be in the range of 60-70%. For WGA, the sensitivity was 65% signifying the fraction of the glycoproteins that were oxidized by the probe. Thus, its utility as a general N-glycan lectin is moderate at least for this cell line. The sensitivity of both SNA and MAL were higher at 68%. Interestingly, the fucose binding lectins, AAL and PSA, were found to be even higher at approximately 72%. The mannose binding protein HHL had the lowest sensitivity at 28% which was due to the low expression of high mannose on PNT2 cells.

Glycoprotein-protein interactions on cell membrane are probed by lectins

A small fraction of the nonglycosylated proteins oxidized by the probes were found to be primarily glycan binding proteins that were oxidized due to their proximity to the glycoproteins. For example, ANXA 2 (Annexin II) is a protein with cationic binding function and was oxidized by SNA. Annexins are a group of calcium-dependent membrane proteins that associate with other proteins. They have been shown to have glycan-binding properties with affinities toward negatively charged glycans such as sialylated glycans and heparan.(30, 31)

SNA and MAL had an 85% overlap in the oxidized glycoprotein targets. The nonglycosylated proteins had a similarly large overlap (over 70%) (**Figure S8a**). We compared the nonglycosylated proteins oxidized by SNA and MAL to those previously identified as potentially sialic acid binding proteins using an orthogonal approach. In an earlier study, sialic acids were linked to an Fe³⁺ probe to mark (oxidize) proteins that were in the proximity of sialic acids.(18) By comparing the current results with the previous, we found more than 60% overlap in the proteins identified further supporting the notion that oxidized nonglycosylated proteins are those that interact with the primary targeted glycoproteins (**Figure S9**).

The nonglycosylated proteins oxidized by fucose-recognizing lectins (AAL and PSA) were similarly believed to be fucose-binding proteins. The overlap in the nonglycosylated proteins between AAL and PSA were over 90% (**Figure S8b**). There was a similarly large overlap (>80%) between fucose-associated and sialic acid-associated proteins (**Figure S10**). The similarities were consistent with the glycosylation in PNT2, which the majority of glycans being both sialylated and fucosylated, the large overlap in the oxidized proteins between the two types of lectins were consistent with the specificities of these interactions.

Other relationships between the target glycoproteins and the associated (nonglycosylated) proteins were further examined using the STRING software. A general map using WGA with Cytoscape is shown (**Figure 5a**), with the glycoproteins in red and nonglycosylated proteins in blue.⁽³²⁾ The interaction map showed that the proteins (glycosylated and non) were highly interactive and mediated by specific types of glycosylation. Similarly, the interaction maps can be generated using other lectins, such as SNA and AAL (**Figure S11**). More than 75% overlap was observed by comparing the SNA and AAL interaction network (**Figure S12a**), which is consistent with the dominant presence of sialofucosylated glycans. For example, a highly sialofucosylated protein EGFR was found to interact extensively with other nonglycosylated proteins and glycoproteins (**Figure 5b**). In contrast, there was less than a 20% overlap between high mannose and sialic acid binding lectins (**Figure S12b**). The results confirmed the consistency of Lectin PROXL and suggested that protein-protein interactions can be probed by various lectins and assigned to the mediating glycan type based on the lectin. These results further suggested that the glycan structure could act as the determinant to control the interactions between glycoproteins.

Application of the method towards LNCaP cell line

The lectin probes were also applied to human prostate cancer cell line, LNCaP, to investigate the behavior of lectins with a cell lines having different glycocalyx profiles (**Supporting Data 5-7**). The LNCaP cell line was found to be highly fucosylated but differs from PNT2 in that it was low in sialylation (**Figure S7**). A smaller subset of the lectins used above was applied to LNCaP. The fucose-binding probe, modified AAL, yielded over 70 unique glycoproteins. The WGA yielded similar numbers reflecting the high level of fucosylation in LNCaP. Increasing the amount of fucosylation on the cell surface in LNCaP, relative to PNT2, increased the number of oxidized glycoproteins by a factor of two. There was similarly a large overlap of oxidized glycoproteins between AAL and WGA. The SNA probe oxidized only 22 glycoproteins due to the lower sialylation of LNCaP (**Figure S13**).

Inspection of glycoproteins were similarly consistent with the specificity of the lectins (**Figure S14a**). For example, glycoproteomic analysis of TFR1 (transferrin receptor protein 1) yielded primarily fucosylated glycans at N251. Both AAL and WGA produced oxidation of TFR1 in close proximity to the glycosylation site at M283. However, SNA did not produce oxidized peptides for TFR1. The sensitivity of the lectin was also determined for the cell line. Same observations were noticed by investigating the sensitivity of the lectins in LNCaP cell line, in which all three lectins yielded more than 60% sensitivity towards the putative targeted proteins (**Figure S14b**).

A comparison between LNCaP (a prostate cancer cell line) and PNT2 (a nonmalignant prostate cell line) was useful for comparing relationships in the protein network. For example, a nonglycosylated protein ACSL1 (long-chain-fatty-acid--CoA ligase 1) and a glycoprotein PPT1

(palmitoyl-protein thioesterase 1) were oxidized by the WGA lectin in LNCaP, however neither were observed in PNT2. Indeed, it has been reported that both ACSL1 and PPT1 were upregulated in prostate cancer cells.(33) STRING analysis showed that ACSL1 do interact with PPT1, and studies have shown that PPT1 glycosylation can affect its ability to form complexes.(34) These results do point to a correlation between ACSL1 and PPT1 in prostate cancer that was potentially mediated by glycosylation.

DISCUSSION

Lectin probes that can label protein targets provide the opportunity to understand the relationships between the lectin, the glycan and polypeptide scaffold. Additionally, they elucidate the environment of the glycoprotein targets by providing the proteins that associate with specific glycoproteins on the cell membrane. Probing lectin specificities purely on the glycan provides an incomplete picture of the lectin-glycan interactions.⁽³⁵⁾ The polypeptide backbone plays a significant role in defining the conformation of glycans. However, isolating the protein masks the effects of cellular conditions such as interaction with other molecules/ions due to specific localization.⁽³⁶⁾ Indeed, the protein backbone is known to restrict glycan conformation and subcellular localization.⁽³⁷⁾ Although lectins are typically used to characterize glycans on cell membrane, this study adds to the very limited research performed *in vitro* to examine the lectin specificity.

The fraction of the putative glycoprotein targets that were oxidatively labeled yields the sensitivity of the lectins. When all lectins were used, the total number of glycoproteins oxidized by the lectin probes corresponded to approximately 70% of all glycoproteins detected with both PNT2 and LNCaP. Although there was a broad diversity in the lectins used in this study, not all glycoproteins were oxidatively labeled by the lectin probes. Unmarked glycoproteins, in general, resulted from at least two reasons, namely the expression level of specific glycans on glycoproteins were low or there were static and dynamic variations in glycoprotein structures.

The results further demonstrated that specific structural motifs such as linkages can be determined at the glycoprotein level by the Lectin PROXL. SNA and MAL are both sialic acid binding lectins with specificities for $\alpha(2,6)$ and $\alpha(2,3)$ sialylated glycans, respectively. More

oxidized proteins were obtained with the SNA probe, indicating a higher expression of $\alpha(2,6)$ sialic acid in PNT2 cell line. Conversely for fucosylated glycans, PSA with a specificity for core $\alpha(1,6)$ fucose yielded less labeled glycoproteins than AAL with broader specificity. Most fucosylated glycans are generally core fucosylated first, followed by antenna fucosylation. That PSA yielded much less oxidation was due more to the shielding of the core fucose by the polypeptide. The poorest specificities were found for the mannose-binding lectin HHL. In contrast, PHA-E and PHA-L which recognizes galactose residues on N-glycans had higher specificities and oxidized over 70% of galactose-containing glycoproteins. WGA with broad specificity for N-glycans also labeled a larger fraction of the glycoproteins. Nonetheless, it too had a unique, previously unreported specificity as it appeared to favor hybrid-type over complex-type glycans.

This Lectin PROXL method also revealed cell surface networks that were mediated by specific glycosylation. The nonglycosylated proteins oxidized by the probes were consistent more with glycan-binding proteins than random, nonspecific lectin interactions. By constructing the interaction networks associated with the lectins, we noticed several glycoproteins that behaved as hubs by simultaneously interacting with several other proteins. For example, EGFR was found to interact with many other glycoproteins and nonglycosylated proteins. Indeed, EGFR plays a central role in many biological processes and associates with many diseases.(38) Thus, along with EGFR, the nonglycosylated proteins catenins such as CTNA1 (catenin alpha-1), CTNB1 (catenin beta-1), and CTND1 (catenin delta-1) were oxidized by the SNA and AAL probes. The interactions between EGFR and catenins have specifically been shown to rely on the glycosylation of EGFR.(39-41) Other known interactions of glycoproteins interacting with other glycoproteins were also obtained in these interaction maps. For example, EGFR and the glycoprotein ITGB4

(integrin beta-4) were both oxidized by SNA. Here too, N-glycans on EGFR were reported to mediate the association between the two glycoproteins.(42) Other associating proteins were also found to be potentially mediated by glycans. For example, the glycoprotein ENPL (endoplasmin) was oxidized by HHL, while CAV1 (Caveolin-1), a nonglycosylated protein, was also oxidized. Examination of the HHL proteins by STRING predicted that both CAV1 and ENPL are strongly interacting proteins. Comparison of the proteins marked HHL with other lectins, for example SNA and AAL, did not yield the same glycoprotein-protein interaction map, suggesting that the interaction map may be mediated by high mannose glycosylation, rather than either sialylated or fucosylated glycans. This result therefore suggested that the interactions between glycoprotein and nonglycosylated proteins depended on the glycan structures, perhaps as expected, but now more specifically elucidated. High mannose glycosylation on the cell membrane is important and have been found to play a role in the migration and invasion of the cells by strengthening extracellular protein complexes.(43)

Lectin PROXL is a new addition to the glycobiology toolbox that reveals the blind spot that limits traditional lectin-based analysis. It identifies the protein scaffold of the glycans as well as the associating proteins in the complex. More specifically in also yields glycan composition and the site-specific localization. As aberrant glycosylation is a hallmark of many diseases including cancer, it will be valuable in developing new targets and new therapeutics. Moreover, the method is not limited to lectins. Future publications will undoubtedly widen the utility of Lectin PROXL to include antibodies and other proteins whose targets on tissues and cell membranes are highly desirable.

ACKNOWLEDGMENTS

Research reported in this report was supported by General Medicine of the National Institutes of Health under the award numbers RO1GM049077 and RO1 AG062240.

AUTHOR CONTRIBUTIONS

Y.X. designed and performed experiments, analyzed data, created the figure and wrote the manuscript. Y.S. performed experiments and analyzed data. Q.L, S.J., and J.R, performed experiments. C.B.L conceived the idea, supervised the study, and co-wrote the manuscript.

CONFLICTS OF INTEREST

There are no conflicts with this report.

FIGURES

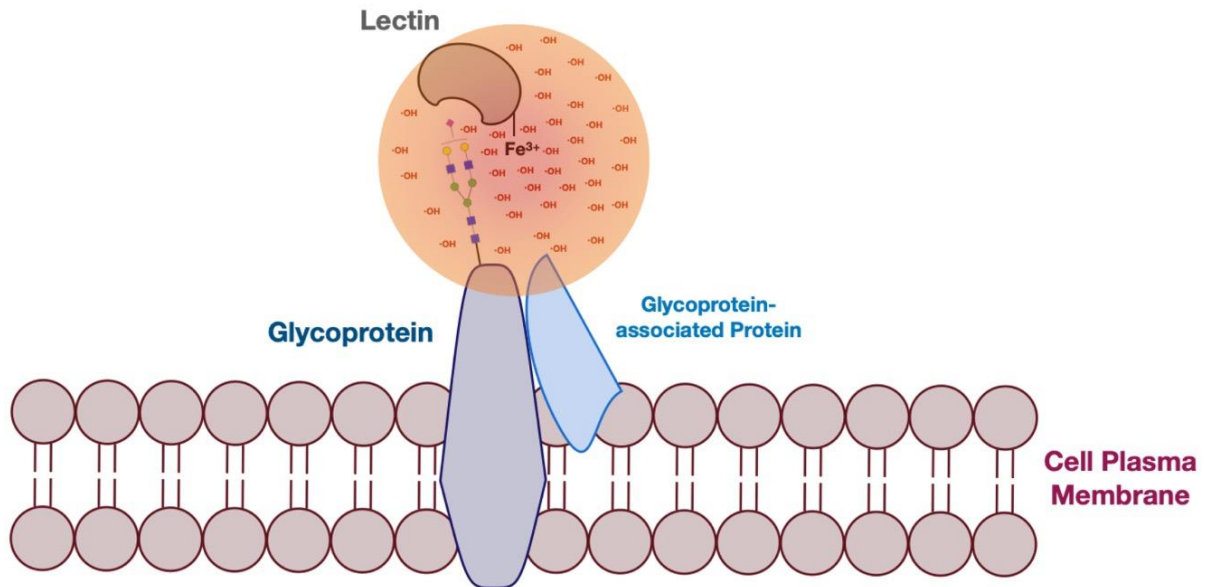


Figure 1. The schematic representation of the determination of the glycoprotein specificity of lectins on the cell membrane through oxidative proteomics. The oxidation probe (Fe(III))-modified lectins were treated to cells, and the hydroxyl radicals were induced by treating cells with hydrogen peroxide nearby the oxidation probe. The localized hydroxyl radicals yielded the oxidation of lectin-targeted glycoproteins and glycoprotein-associated proteins.

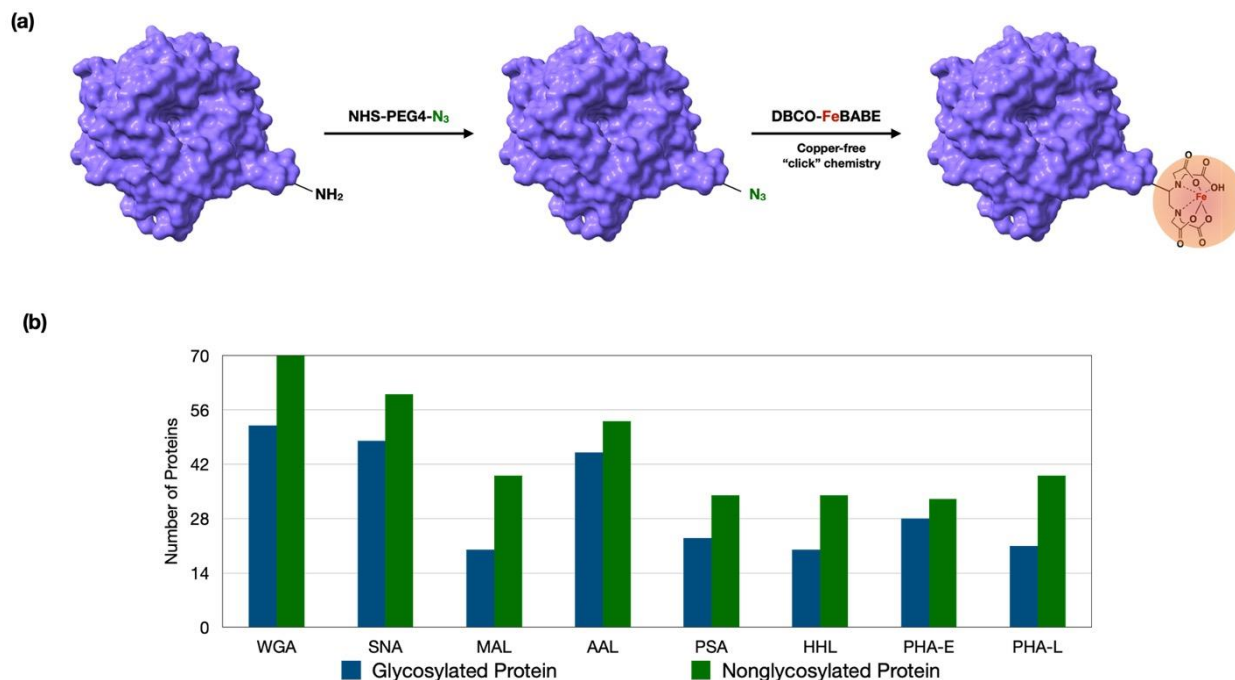


Figure 2. (a) The schematic representation of the modification of lectins with Fe(III) probe. The modification was involved in two steps. First, azido group was introduced to the lectin through reacting with a primary amine on lysine side chains, followed by conjugation of synthesized DBCO-FeBABA to the lectin *via* copper-free “click” chemistry. (b) The number of oxidized glycoproteins and non-glycosylated proteins on PNT2 cell line generated under the optimal conditions.

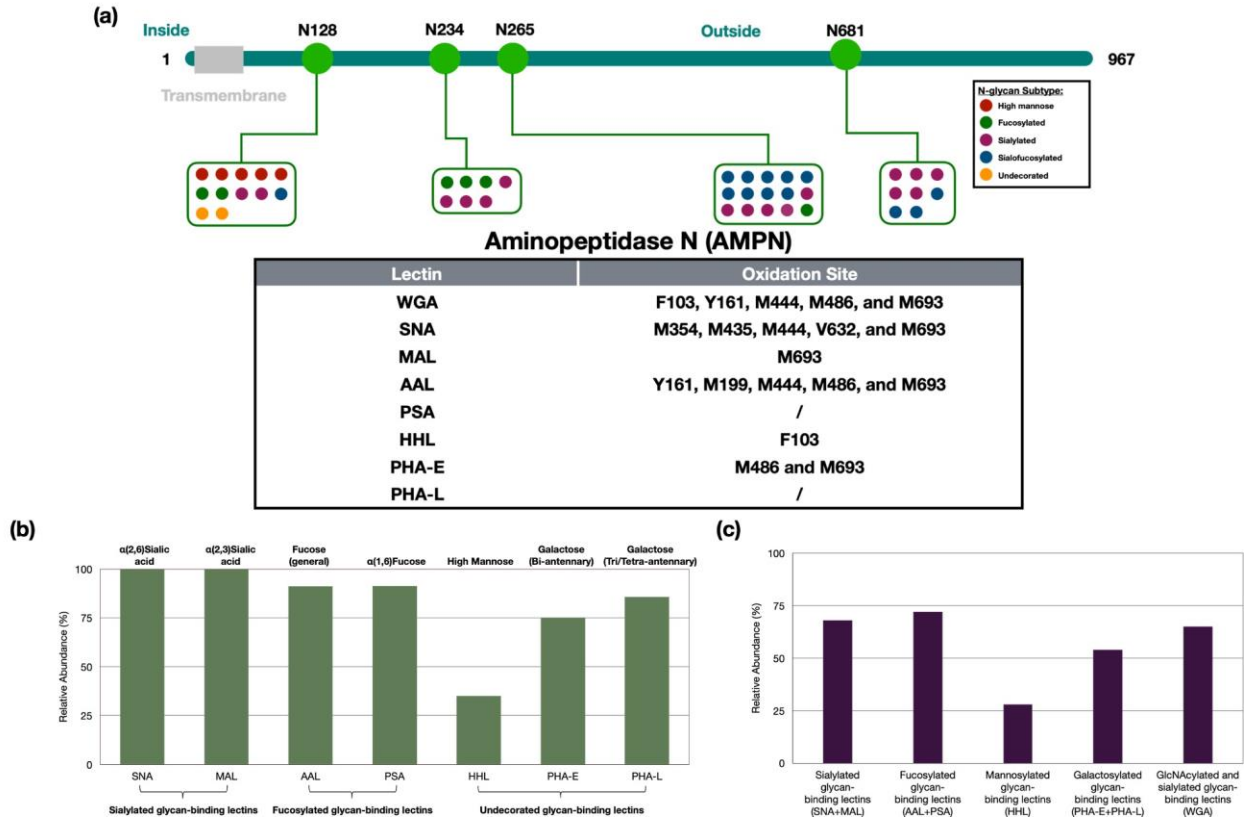
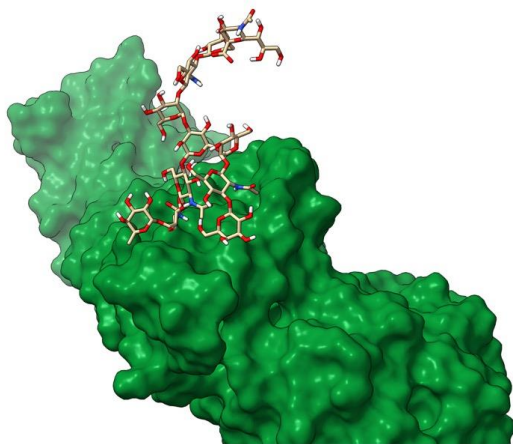
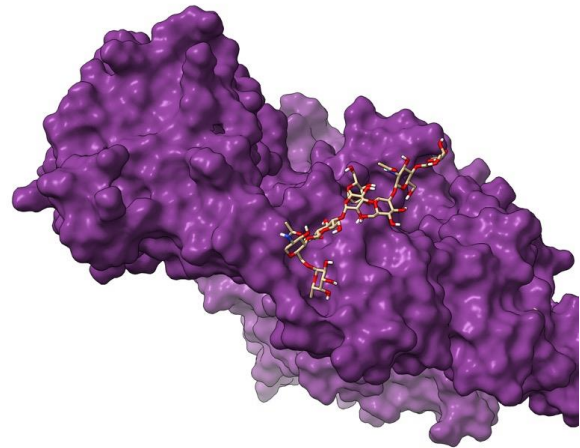


Figure 3. (a) The relationship between sites of glycosylation and sites of oxidation. From the oxidation results, the oxidation site on glycoprotein AMPN (Aminopeptidase N) was highly dependent on the distribution of different types of glycans. (b) The glycoprotein specificity of lectins on PNT2 cells. Most of the lectins showed the high specificity towards the targeted glycoproteins (>70%), while HHL did not recognize the high mannose N-glycans exclusively. (c) The sensitivity of the method on PNT2 cells. Most of the sensitivity was calculated to be in the range of 60-70%.



ITGB1



EGFR

Figure 4. The three-dimensional structures of glycoproteins ITGB1 (integrin beta-1, PDB: 3VI3) and EGFR (epidermal growth factor receptor, PDB: 1NQL) containing the glycan $\text{Man}_{(3)}\text{Gal}_{(2)}\text{GlcNAc}_{(4)}\text{Fuc}_{(1)}\text{Sia}_{(1)}$. The glycoprotein models were built using Glycam (<http://glycam.org>).

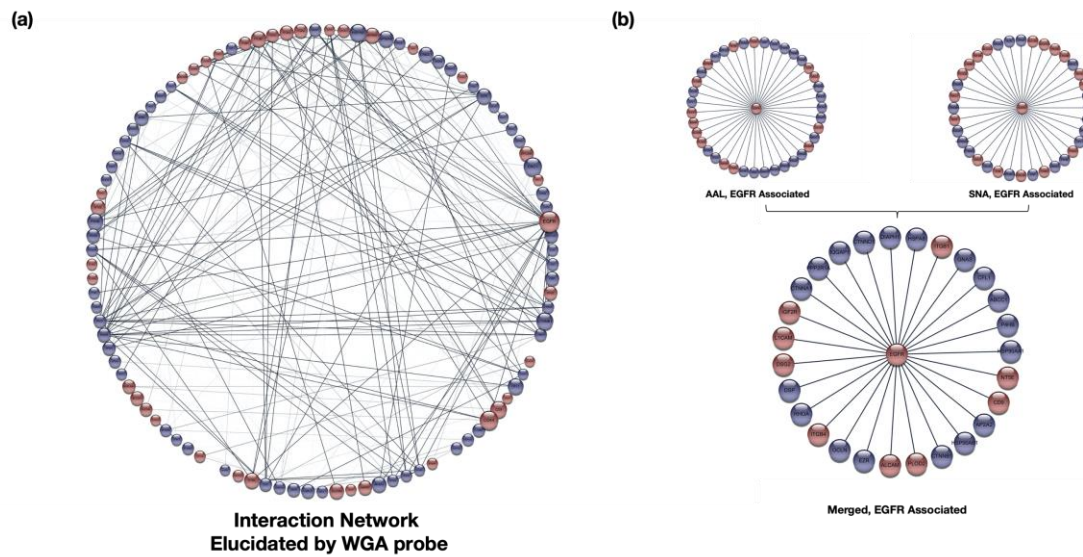


Figure 5. (a) The WGA interaction network was analyzed using STRING software (<https://string-db.org>).⁽⁴⁴⁾ The glycoproteins were colored in red and non-glycosylated proteins were colored in blue. The size of the node corresponded to the number of interactions involved. The weight of each protein connection showed the confidence in the interactions. (b) The interaction network among EGFR-associated proteins was revealed by AAL and SNA probes. Over 80% overlap was observed in the two interaction networks.

TABLES

Lectin	Origin	Monosaccharide Specificity	Putative N-Glycan Specificity
WGA	<i>Wheat germ</i>	N-Acetylglucosamine and sialic acid	N-Glycans in general
SNA	<i>Sambucus nigra</i>	Sialic Acid	$\alpha(2,6)$ Sialylated glycans
MAL	<i>Maackia amurensis</i>	Sialic Acid	$\alpha(2,3)$ Sialylated glycans
AAL	<i>Aleuria aurantia</i>	Fucose	Fucosylated glycans in general
PSA	<i>Pisum sativum</i>	Fucose	$\alpha(1,6)$ Fucosylated glycans
HHL	<i>Hippeastrum hybrid</i>	Mannose	High mannose glycans
PHA-E	<i>Phaseolus vulgaris</i>	Galactose	Biantennary complex-type and bisecting glycans
PHA-L	<i>Phaseolus vulgaris</i>	Galactose	Tri/Tetra-antennary complex-type glycans

Table 1. The lectins included in this method and their specificities towards monosaccharides and N-glycans.

REFERENCE

1. L. R. Ruhaak, G. Xu, Q. Li, E. Goonatileke, C. B. Lebrilla, Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chemical Reviews* **118**, 7886-7930 (2018).
2. M. J. Paszek *et al.*, The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature* **511**, 319-325 (2014).
3. H. J. An, S. R. Kronewitter, M. L. A. de Leoz, C. B. Lebrilla, Glycomics and disease markers. *Current Opinion in Chemical Biology* **13**, 601-607 (2009).
4. M. E. Janik, A. Lityńska, P. Vereecken, Cell migration—The role of integrin glycosylation. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1800**, 545-555 (2010).
5. M. D. Langer, H. Guo, N. Shashikanth, J. M. Pierce, D. E. Leckband, N-glycosylation alters cadherin-mediated intercellular binding kinetics. *Journal of Cell Science* **125**, 2478 (2012).
6. D. Park *et al.*, Enterocyte glycosylation is responsive to changes in extracellular conditions: implications for membrane functions. *Glycobiology* **27**, 847-860 (2017).
7. B. Belardi, Carolyn R. Bertozzi, Chemical Lectinology: Tools for Probing the Ligands and Dynamics of Mammalian Lectins In Vivo. *Chemistry & Biology* **22**, 983-993 (2015).
8. Y. Xie *et al.*, Mannose-based graft polyesters with tunable binding affinity to concanavalin A. *Journal of Polymer Science Part A: Polymer Chemistry* **55**, 3908-3917 (2017).
9. K. Godula *et al.*, Control of the Molecular Orientation of Membrane-Anchored Biomimetic Glycopolymers. *Journal of the American Chemical Society* **131**, 10263-10268 (2009).
10. T. Tanaka *et al.*, Protecting-Group-Free Synthesis of Glycopolymers Bearing Sialyloligosaccharide and Their High Binding with the Influenza Virus. *ACS Macro Letters* **3**, 1074-1078 (2014).
11. Y. Chen, L. Ding, H. Ju, In Situ Cellular Glycan Analysis. *Accounts of Chemical Research* **51**, 890-899 (2018).
12. J. Han *et al.*, Laser cleavable probes for in situ multiplexed glycan detection by single cell mass spectrometry. *Chemical Science* **10**, 10958-10962 (2019).
13. S. Han, B. E. Collins, P. Bengtson, J. C. Paulson, Homomultimeric complexes of CD22 in B cells revealed by protein-glycan cross-linking. *Nature Chemical Biology* **1**, 93-97 (2005).
14. T. N. C. Ramya *et al.*, &em>In Situ trans&/em> Ligands of CD22 Identified by Glycan-Protein Photocross-linking-enabled Proteomics. *Molecular & Cellular Proteomics* **9**, 1339 (2010).
15. Y. Tanaka, J. J. Kohler, Photoactivatable Crosslinking Sugars for Capturing Glycoprotein Interactions. *Journal of the American Chemical Society* **130**, 3278-3279 (2008).
16. S. A. Datwyler, C. F. Meares, Protein–protein interactions mapped by artificial proteases: where σ factors bind to RNA polymerase. *Trends in Biochemical Sciences* **25**, 408-414 (2000).
17. X. R. Liu, M. M. Zhang, M. L. Gross, Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chemical Reviews*, (2020).

18. Q. Li, Y. Xie, G. Xu, C. B. Lebrilla, Identification of potential sialic acid binding proteins on cell membranes by proximity chemical labeling. *Chemical Science* **10**, 6199-6209 (2019).
19. N. Shibuya *et al.*, The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence. *Journal of Biological Chemistry* **262**, 1596-1601 (1987).
20. W. C. Wang, R. D. Cummings, The immobilized leukoagglutinin from the seeds of *Maackia amurensis* binds with high affinity to complex-type Asn-linked oligosaccharides containing terminal sialic acid-linked alpha-2,3 to penultimate galactose residues. *Journal of Biological Chemistry* **263**, 4576-4585 (1988).
21. H. Tatenno, S. Nakamura-Tsuruta, J. Hirabayashi, Comparative analysis of core-fucose-binding lectins from *Lens culinaris* and *Pisum sativum* using frontal affinity chromatography. *Glycobiology* **19**, 527-536 (2009).
22. S. Hammarström, M. L. Hammarström, G. Sundblad, J. Arnarp, J. Lönngrén, Mitogenic leukoagglutinin from *Phaseolus vulgaris* binds to a pentasaccharide unit in N-acetyllactosamine-type glycoprotein glycans. *Proceedings of the National Academy of Sciences* **79**, 1611 (1982).
23. T. Irimura, T. Tsuji, S. Tagami, K. Yamamoto, T. Osawa, Structure of a complex-type sugar chain of human glycoporphin A. *Biochemistry* **20**, 560-566 (1981).
24. H. Kaku, E. J. M. Van Damme, W. J. Peumans, I. J. Goldstein, Carbohydrate-binding specificity of the daffodil (*Narcissus pseudonarcissus*) and amaryllis (*Hippeastrum hybr.*) bulb lectins. *Archives of Biochemistry and Biophysics* **279**, 298-304 (1990).
25. A. Yamamoto, R. Masaki, Y. Tashiro, Characterization of the isolation membranes and the limiting membranes of autophagosomes in rat hepatocytes by lectin cytochemistry. *Journal of Histochemistry & Cytochemistry* **38**, 573-580 (1990).
26. A. C. Mello Filho, R. Meneghini, In vivo formation of single-strand breaks in DNA by hydrogen peroxide is mediated by the Haber-Weiss reaction. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **781**, 56-63 (1984).
27. D. Wu, J. Li, W. B. Struwe, Carol V. Robinson, Probing N-glycoprotein microheterogeneity by lectin affinity purification-mass spectrometry analysis. *Chemical Science* **10**, 5146-5155 (2019).
28. K. N. Kirschner *et al.*, GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *Journal of Computational Chemistry* **29**, 622-655 (2008).
29. J. T. Hirabayashi, H.; Shikanai, T.; Aoki-Kinoshita, K.F.; Narimatsu, H., The Lectin Frontier Database (LfDB), and Data Generation Based on Frontal Affinity Chromatography. *Molecules* **20**, 951-973 (2015).
30. C. Shao *et al.*, Crystallographic Analysis of Calcium-dependent Heparin Binding to Annexin A2. *Journal of Biological Chemistry* **281**, 31689-31695 (2006).
31. M. N. Kundranda *et al.*, The Serum Glycoprotein Fetuin-A Promotes Lewis Lung Carcinoma Tumorigenesis via Adhesive-Dependent and Adhesive-Independent Mechanisms. *Cancer Research* **65**, 499 (2005).
32. P. Shannon *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).

33. Z. He, X. Duan, G. Zeng, Identification of potential biomarkers and pivotal biological pathways for prostate cancer using bioinformatics analysis methods. *PeerJ* **7**, e7872 (2019).
34. A. Lyly *et al.*, Glycosylation, transport, and complex formation of palmitoyl protein thioesterase 1 (PPT1) – distinct characteristics in neurons. *BMC Cell Biology* **8**, 22 (2007).
35. R. A. Dwek, Glycobiology: Toward Understanding the Function of Sugars. *Chemical Reviews* **96**, 683-720 (1996).
36. M. Schaffer *et al.*, Optimized cryo-focused ion beam sample preparation aimed at in situ structural studies of membrane proteins. *Journal of Structural Biology* **197**, 73-82 (2017).
37. D. D. Park *et al.*, Membrane glycomics reveal heterogeneity and quantitative distribution of cell surface sialylation. *Chemical Science* **9**, 6271-6285 (2018).
38. P. Wee, Z. Wang, Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers* **9**, (2017).
39. J. Li *et al.*, Perturbation of the mutated EGFR interactome identifies vulnerabilities and resistance mechanisms. *Molecular Systems Biology* **9**, 705 (2013).
40. M. Azimzadeh Irani, S. Kannan, C. Verma, Role of N-glycosylation in EGFR ectodomain ligand binding. *Proteins: Structure, Function, and Bioinformatics* **85**, 1529-1549 (2017).
41. V. Tajadura-Ortega *et al.*, O-linked mucin-type glycosylation regulates the transcriptional programme downstream of EGFR in breast cancer. *bioRxiv*, 714675 (2019).
42. Y. Kariya, J. Gu, N-Glycosylation of β 4 Integrin Controls the Adhesion and Motility of Keratinocytes. *PLOS ONE* **6**, e27084 (2011).
43. D. D. Park *et al.*, Metastasis of cholangiocarcinoma is promoted by extended high-mannose glycans. *Proceedings of the National Academy of Sciences*, 201916498 (2020).
44. D. Szklarczyk *et al.*, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research* **47**, D607-D613 (2018).

SUPPLEMENTARY INFORMATION

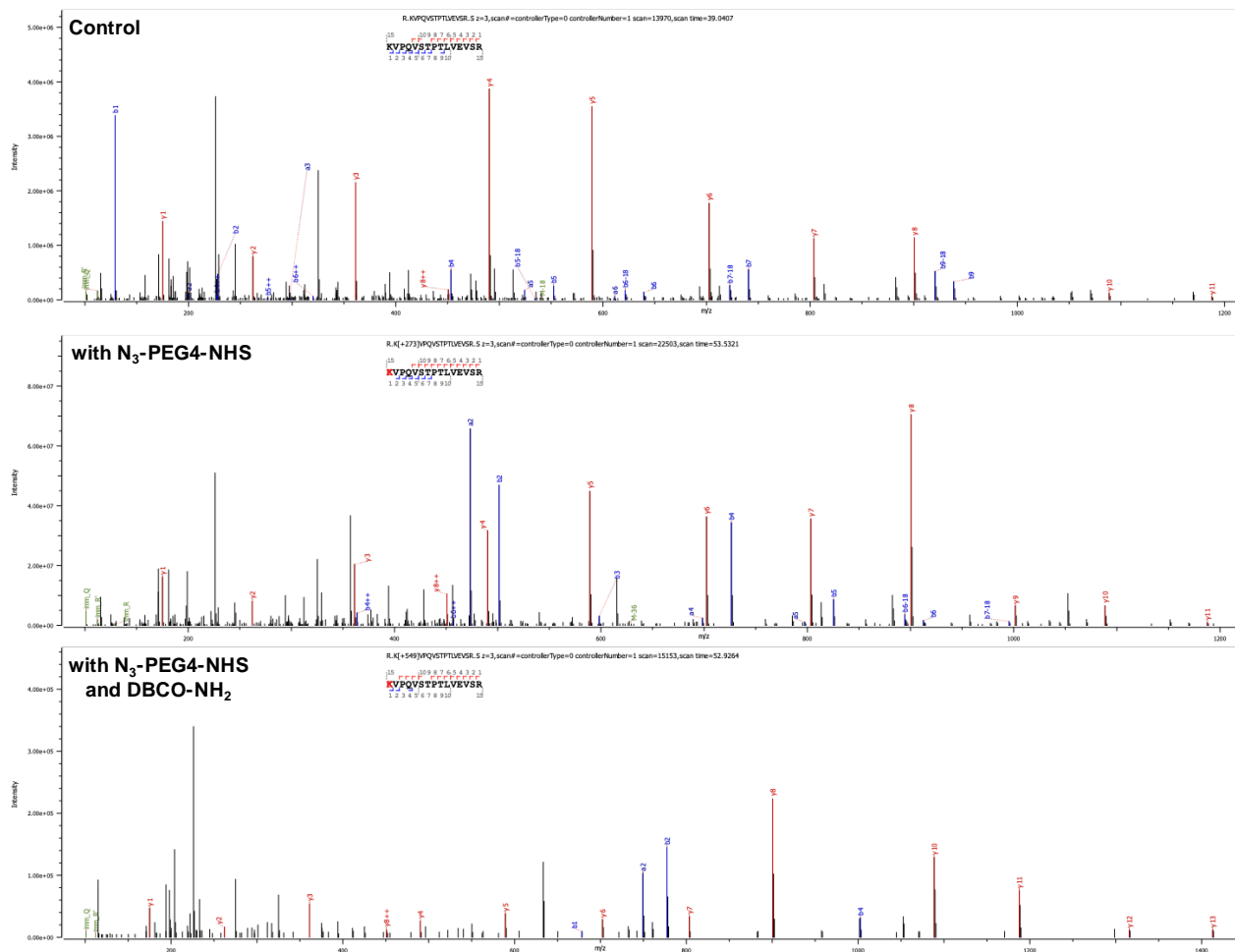


Figure S4.1. The extent of reaction was determined using a model protein, BSA. The product peptides were characterized by nanoLC-MS. The tandem MS/MS data showed the modification with azido group (+273.13 Da) and DBCO (+549.26 Da) at K437 residue.

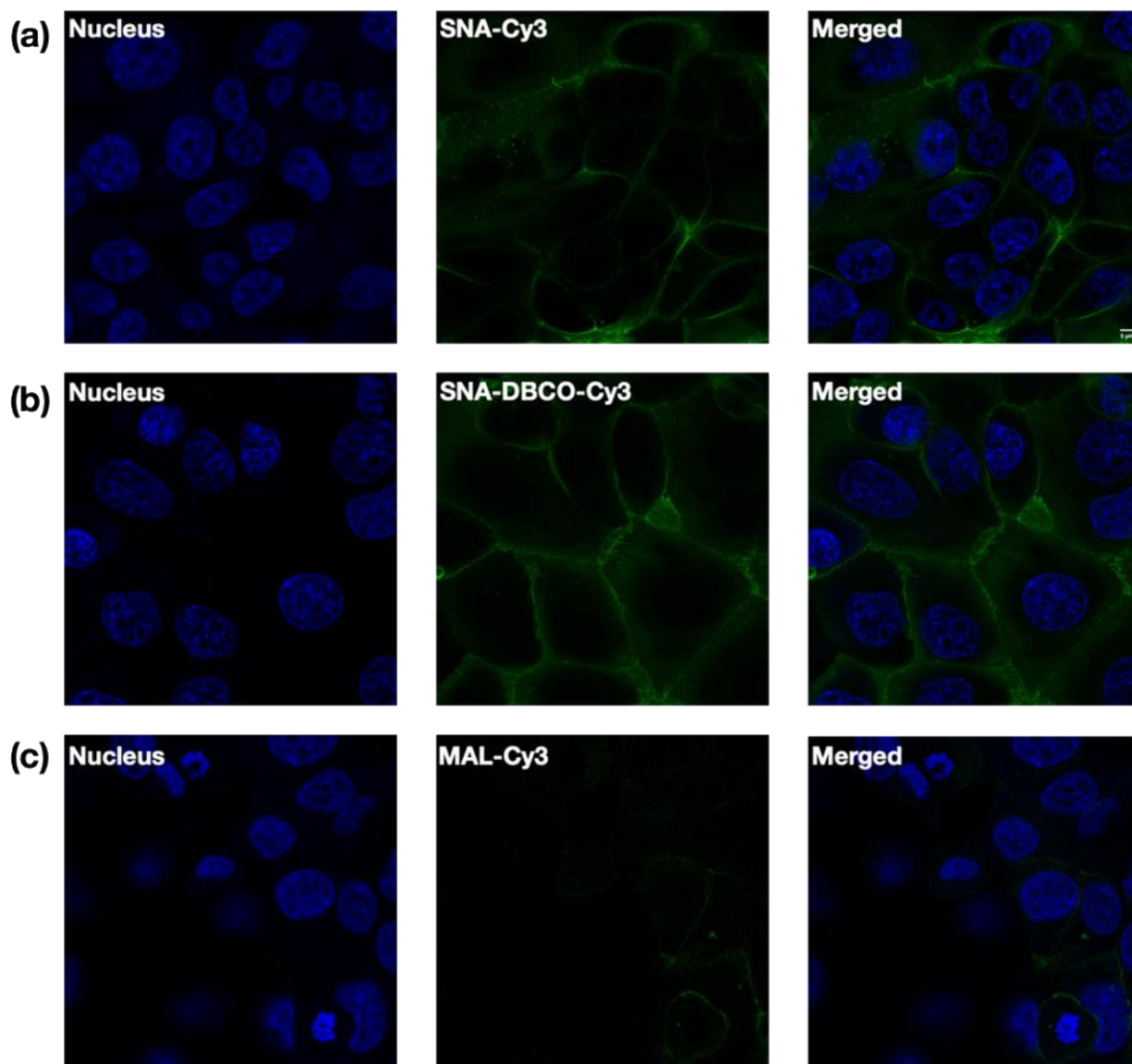


Figure S4.2. Validation of the binding efficiencies of the modified lectins using confocal microscopy. PNT2 Cells were treated with **(a)** SNA-Cyanine3 (control), **(b)** SNA-DBCO-Cyanine3 (modified), and **(c)** MAL-Cyanine3. The cell nucleus was stained with Hoechst 33342.

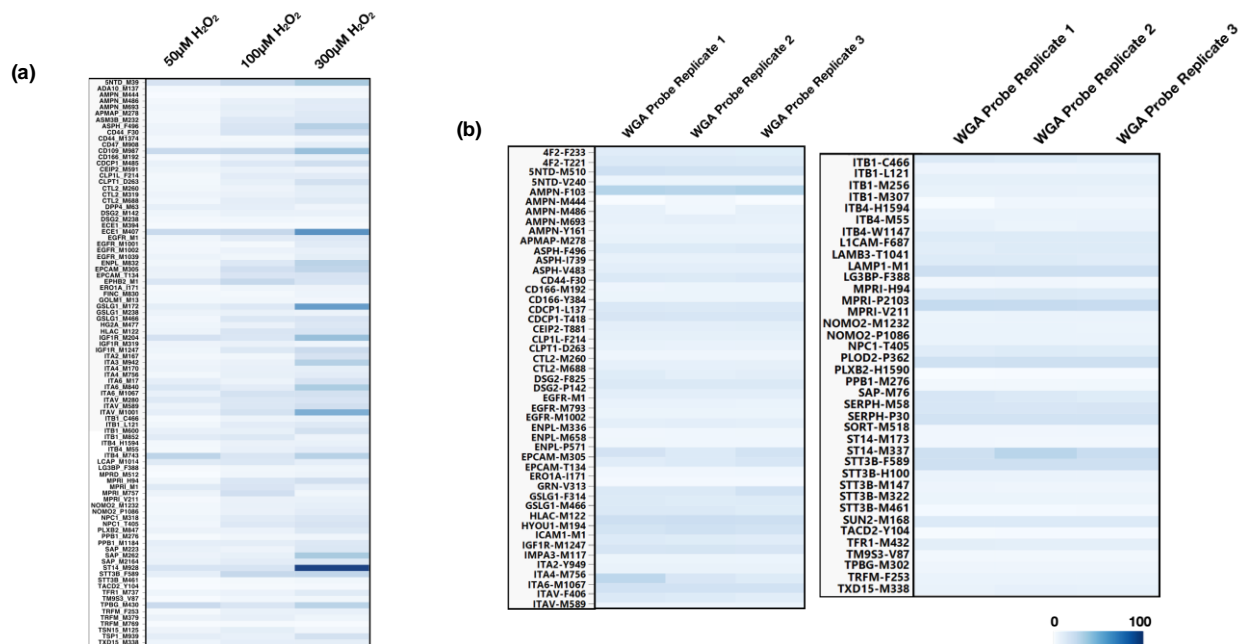


Figure S4.3. (a) Optimization of the conditions for the labeling reaction. The extent of oxidation quantified using Byologic software. Each column represents a single treatment condition, and each row represents the oxidation sites of selected proteins. **(b)** Quantification results from three consecutive LC-MS injections of WGA probe-oxidized proteins.

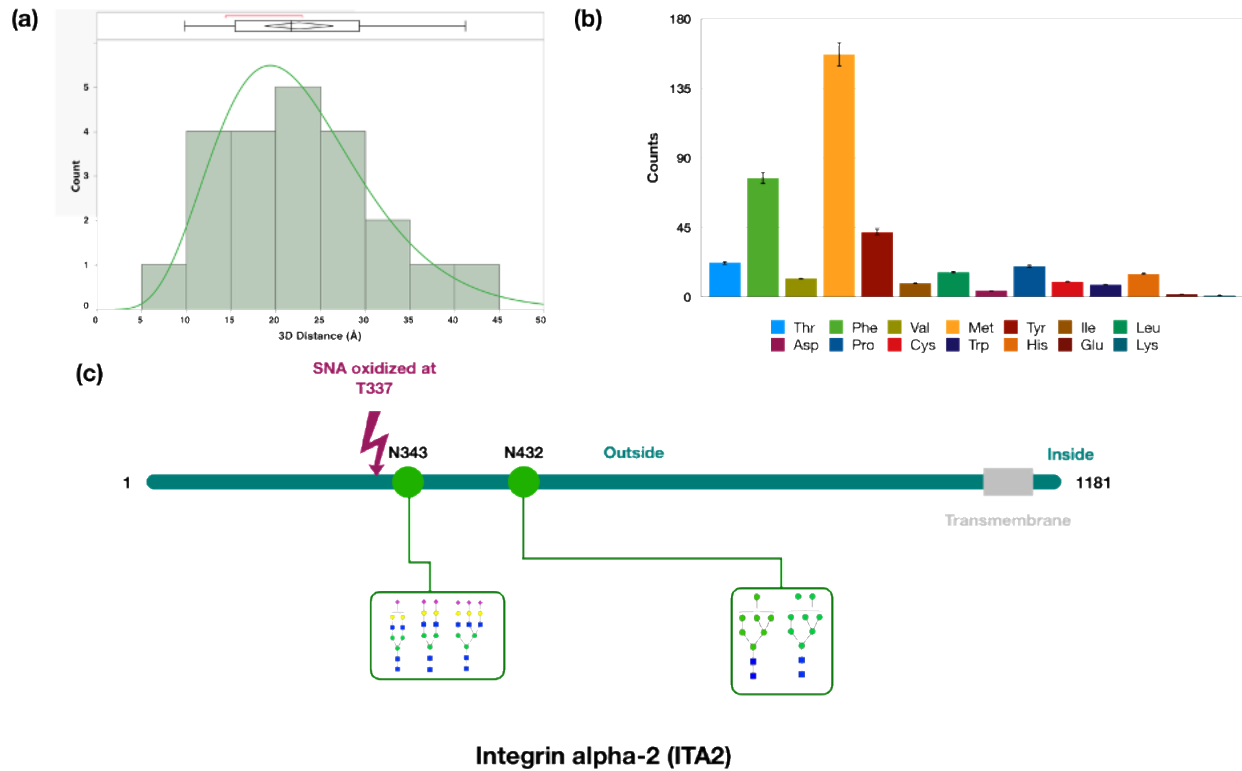


Figure S4.4. (a) The distribution of distances in oxidized glycoproteins between the site of oxidation and glycosylation for WGA. (b) Frequency of oxidized amino acid residues observed from all eight lectin probes. (c) The relationship between glycosylation sites and the site of oxidation on the glycoprotein ITA2 (integrin alpha-2).

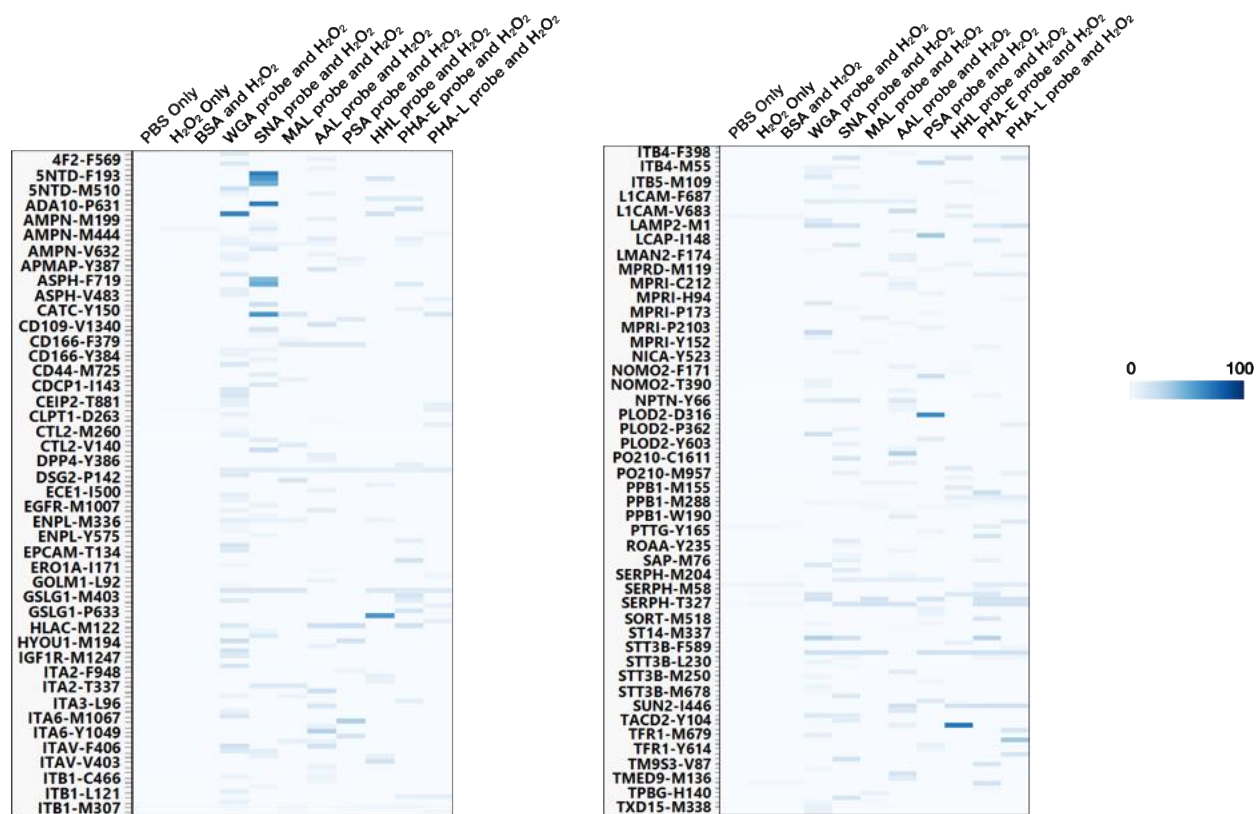


Figure S4.5. The extent of oxidation quantified by Byologic for glycoproteins oxidized by the lectin probes.



Figure S4.6. The overlap in oxidized glycoproteins between different lectins. **(a)** Glycoproteins were found oxidized by both sialylated glycan-binding lectins, SNA and MAL. **(b)** Glycoproteins were found oxidized by both fucosylated glycan-binding lectins, AAL and PSA.

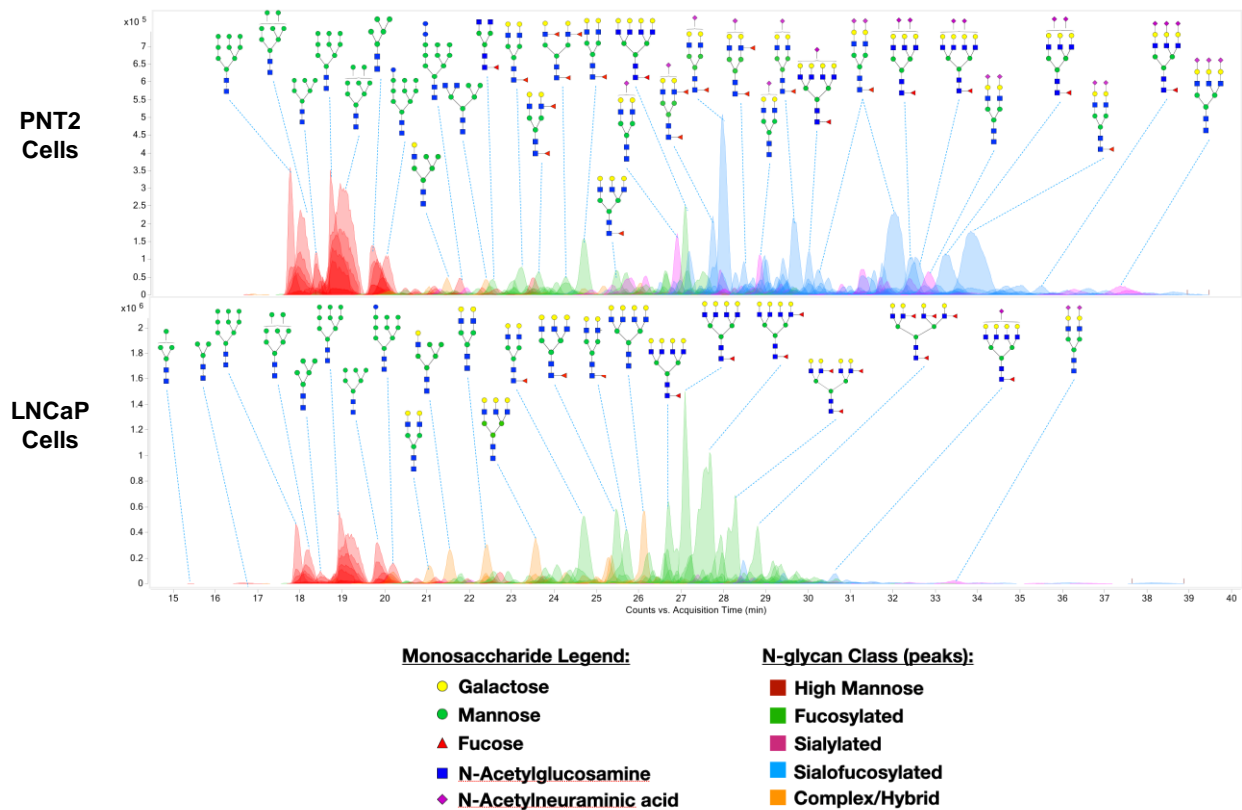


Figure S4.7. LC-MS profile of N-Glycans released from PNT2 cells (top) and LNCaP cells (bottom). Annotated structures are putative based on mass and compositions. LC-MS peaks were color coded to assign glycan subtype.

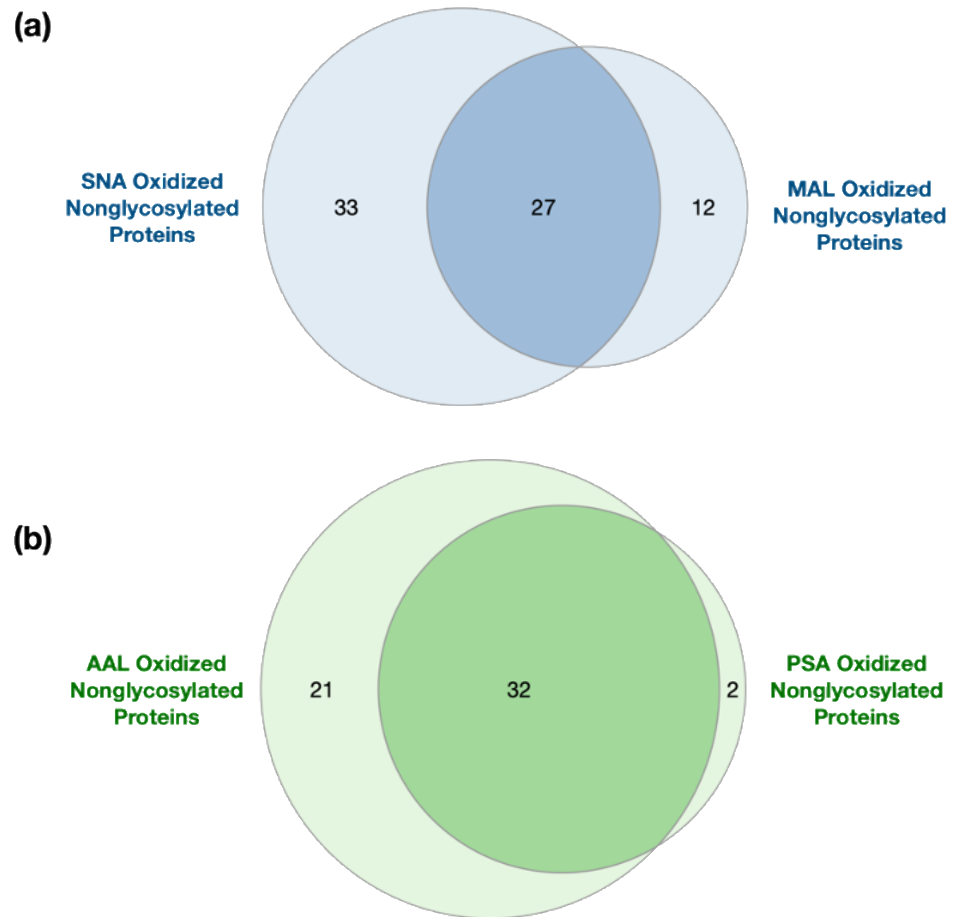


Figure S4.8. The overlap of oxidized nonglycosylated proteins between different lectins.

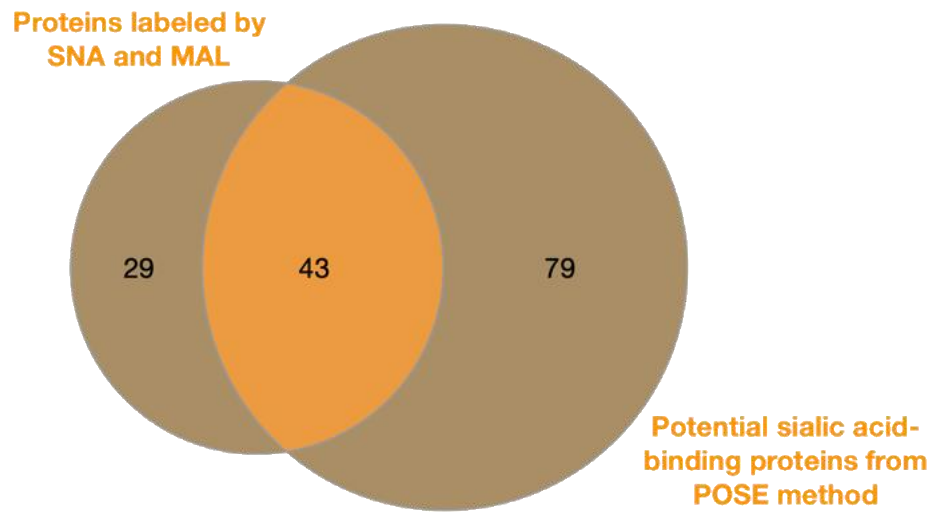


Figure S4.9. Proteins labeled by the sialylated glycan-binding lectins were found to overlap with the potential sialic acid-associated proteins as determined by a previously used method POSE to determine sialic acid binding proteins.

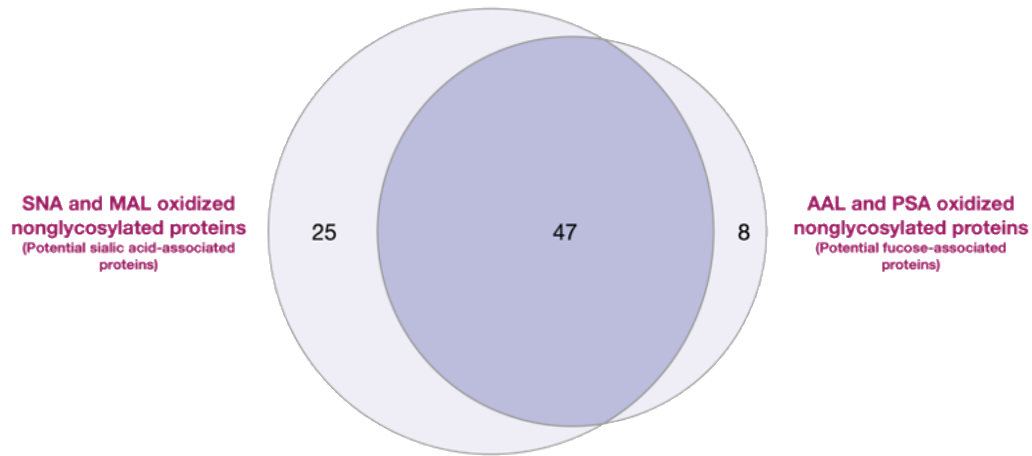
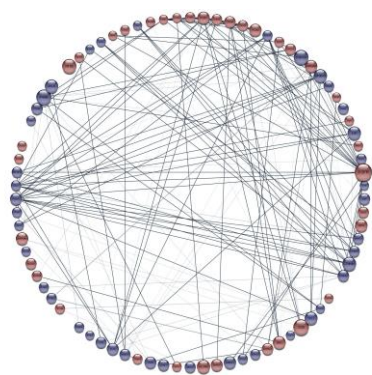
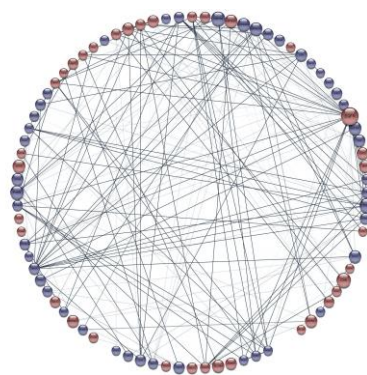


Figure S4.10. Nonglycosylated proteins oxidized by specific lectins. These proteins were oxidized due to their proximity to the respective glycosylated protein. The large overlap supports the observation that show most of the glycans on the cell membrane are both sialylated and fucosylated.



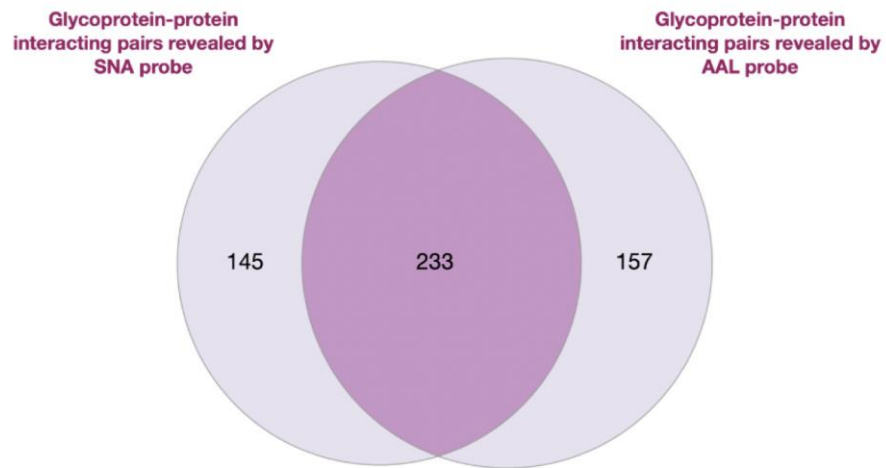
**Interaction Network
from AAL probe**



**Interaction Network
from SNA probe**

Figure S4.11. The interaction networks probed by AAL lectin (left) and SNA lectin (right).

(a)



(b)

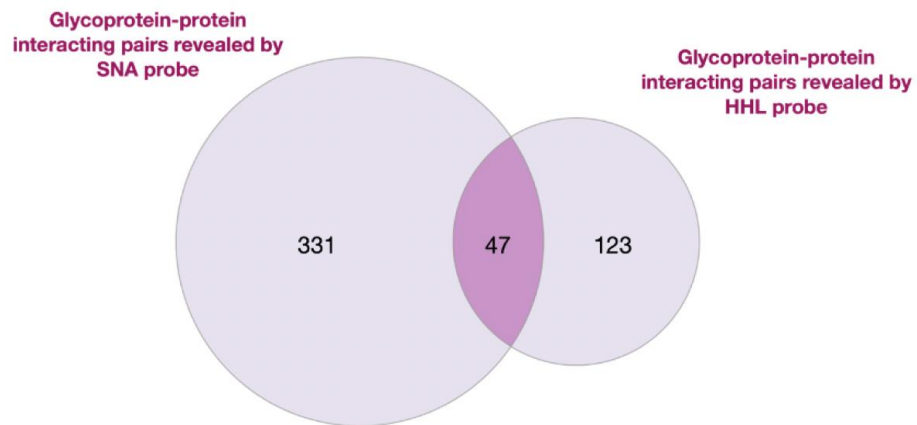


Figure S4.12. (a) The overlapping interaction network found common between SNA and AAL probes. **(b)** The interaction network found common between SNA and HHL probes.

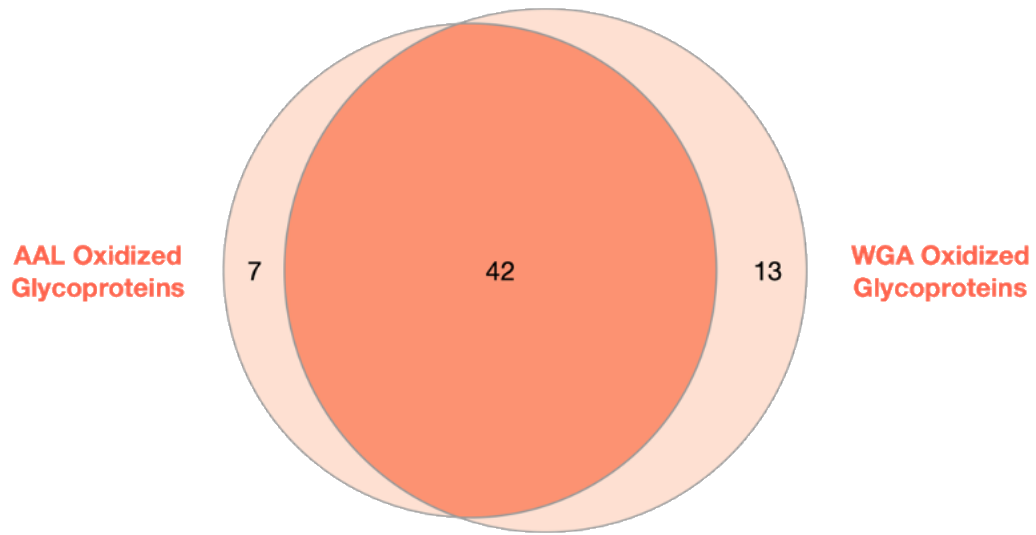


Figure S4.13. The oxidized glycoproteins overlap between lectin WGA and lectin AAL from LNCaP cell line.

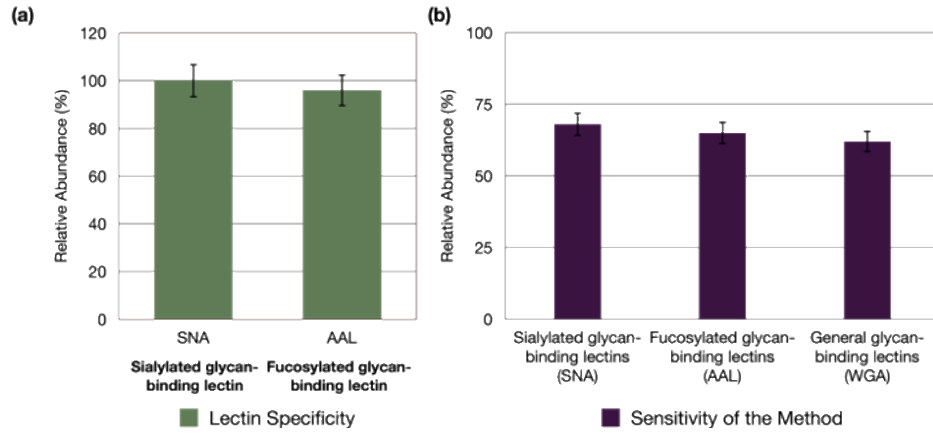


Figure S4.14. (a) The glycoprotein specificity of lectins on LNCaP cells. **(b)** The sensitivity of the method on LNCaP cells. The error bars were obtained based on triplicate results.