

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Investigating the Glycosylation of Immune Cells and Immunoglobulins Through Liquid Chromatography-Mass Spectrometry Methods

Permalink

<https://escholarship.org/uc/item/0zc424dq>

Author

Oloumi, Armin

Publication Date

2024

Peer reviewed|Thesis/dissertation

Investigating the Glycosylation of Immune Cells and Immunoglobulins Through Liquid
Chromatography-Mass Spectrometry Methods

By

ARMIN OLOUMI
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

CARLITO B. LEBRILLA, Chair

SHEILA DAVID

MATTHEW AUGUSTINE

Committee in Charge

2024

ACKNOWLEDGEMENTS

I would like to thank my family for providing me the opportunity to pursue my Ph.D. degree. I would also like to thank my grandparents for their unconditional love and support throughout my life and Ph.D. journey.

I am incredibly thankful to Dr. Carlito B. Lebrilla for accepting me into his research group and mentoring me throughout my Ph.D. education. I would like to thank the Lebrilla laboratory alumni: Dr. Dayoung Park, Dr. Maurice Wong, Dr. Axe Xie, Dr. Ying Sheng, Dr. Qing (Dave) Zhou, Dr. Jennyfer Tena, Dr. Juan Castillo, Dr. Siyu (Cathy) Chen, Dr. Anita Vinjamuri, Dr. Garret Couture, Dr. Nikita (Nikko) Bacalzo, Dr. Ryan Schindler, Diane Tu and Ye (Winnie) Chen. Additionally, I am grateful to the current Lebrilla group members: Sheryl Joyce Grijaldo-Alvarez, Michael Russelle Alvarez, Yasmine Bouchibti, Cheng-Yu (Charlie) Weng, Aaron Stacy, Xavier Holmes, Yean (Shawn) Cheang, Chris Suarez, Sophia Jiang, Yiyun Liu and Riya Gogte for their help and support. I am also incredibly thankful to the undergraduate students, Diego Alvarez and Jonah Hostettler, who worked with me through my Ph.D. journey.

I am grateful for having the opportunity to collaborate with Dr. Emanuel Maverakis, Dr. Stephanie T. Lee, Dr. Angela Zivkovic, Dr. Lee-Way Jin, Dr. Izumi Maezawa, Dr. Daniela Harvey, Dr. Mariana Barboza, Dr. Justin Siegel, Dr. Somen Nandi, Dr. Karen McDonald, Dr. Roland Faller, Dr. Xi Chen, Dr. Abhaya Dandekar and Dr. Priya Shah.

I would like to thank my dissertation committee members, Dr. Sheila David, and Dr. Matt Augustine, for their help and guidance with my qualifying exam, third-year seminar and dissertation. Furthermore, I am grateful to the National Institutes of Health (R01GM049077, R01AG062240), and Dairy Management Inc. for providing funding for the research projects I worked in.

Abstract

There have been significant improvements in characterizing glycoproteins and glycans in the recent years. This dissertation focuses on further expanding the tools for glycan and glycoprotein characterizations and investigating the biological roles of glycans on immune cells and antibodies. Chapter I introduces glycans with focus on their biosynthesis, structures, and biological functions especially in the context of immunity and immunoglobulins. Additionally, the analytical tools for studying glycans and glycoproteins are presented.

Chapter II uses analytical methods to unravel changes in glycosylation of immune cell membrane and their secreted antibodies (IgG1) with glycoenzyme inhibitors, and monosaccharide supplementations. Furthermore, the interactions of IgG1 glycans with Fc receptors was explored through molecular modeling. Effects of glycan changes on cell transcriptome and viability were also assessed by 3' Tag-RNA-Seq and cell proliferation assay. This is the first study that explores changes in glycosylation of cell membrane and secreted proteins in parallel and evaluates changes in transcriptome as the result of induced glycan changes with glycoenzyme inhibitors.

Chapter III probes deeper into the context of glycosylation changes and their effects on immunity and diseases by identifying antibody glycan biomarkers in a rare autoimmune disease (n=10, pemphigus vulgaris). Additionally, the enrichment of glycan markers on auto-reactive antibodies is demonstrated by capturing the disease specific autoantibodies on Desmoglein-3

(Dsg3) target protein using ELISA, and glycoproteome profiling of the captured antibodies using LC-MS/MS.

Chapter IV presents a novel approach for absolute quantitation of proteins, which addresses the issues often encountered in determining protein abundances, namely the lack of protein standards and heterogeneity of PTMs in proteins. This method combines qualitative analysis (protein identification) and quantification (absolute concentration measurements) of proteins into a single analysis. The method was applied to a set of bovine milk protein samples (n = 60) in duplicates, yielding the absolute concentrations of over 600 proteins in bovine milk.

ABBREVIATIONS

2-F-Fucose: 2-deoxy-2-fluorofucose

3-F-Sia: 2,4,7,8,9-Penta-O-acetyl-N-acetyl-3-fluoro-b-d-neuraminic acid

ACPA: Anti-Citrullinated Protein Antibody

ADCC: Antibody-Dependent Cellular Cytotoxicity

BCA: Bicinchoninic Acid

CID: Collision Induced Dissociation

CMAH: Cytidine Monophospho-N-acetylneuraminic Acid Hydroxylase

CMP: Cytidine Monophosphate

CDC: Complement-Dependent Cytotoxicity

DMRM: Dynamic Multiple Reaction Monitoring

DPBS: Dulbecco's Phosphate-Buffered Saline

DTT: Dithiothreitol

ESI: Electrospray Ionization

ELISA: Enzyme-Linked Immunosorbent Assay

FT: Fourier Transform

Fuc: Fucose

Gal: Galactose

GalNAc: N-Acetyl-D-galactosamine

GDP: Guanosine Diphosphate

GC: Gas Chromatography

Glc: Glucose

GlcNAc: N-Acetyl-D-glucosamine

GLUT: Glucose Transporters

HCD: High-Energy Collision Induced Dissociation

Hex: Hexose

HexNAc: N-Acetyl-hexosamine

HILIC: Hydrophilic Interaction Liquid Chromatography

IgA: Immunoglobulin A

IgD: Immunoglobulin D

IgE: Immunoglobulin D

IgG: Immunoglobulin G

IgM: Immunoglobulin M

IAA: Indole-3-Acetic Acid

IIF: Indirect Immunofluorescence

Kif: Kifunensine

LC: Liquid Chromatography

Man: Mannose

MRM: Multiple Reaction Monitoring

MS: Mass Spectrometry

MS/MS: Tandem Mass Spectrometry

Nano-LC: Nanoflow-Liquid Chromatography

NeuAc: N-Acetylneuraminic Acid

NeuGc: N-Glycolylneuraminic acid

PBC: Primary Biliary Cirrhosis

PPM: Parts Per Million

PGC: Porous Graphitic Carbon

PNGase F: Peptide N-Glycosidase F

PP: Total Protein Concentration

PTM: Post Translational Modification

PV: Pemphigus Vulgaris

RPA: Relative Protein Abundance

SGLUT: Sodium-Dependent Glucose Transporters

SPE: Solid-Phase Extraction

TPC: Total Protein Concentration

TOF: Time-of-Flight

UDP: Uridine Diphosphate

UHPLC: Ultra-High Performance Liquid Chromatograph

Table of Contents

Acknowledgements.....	ii
Abstract.....	iv
Abbreviations	vi
Table of Contents	ix

Chapter 1. Introduction to Glycans, Immunoglobulins and Liquid Chromatography-Mass Spectrometry Methods for Glycomics, Glycoproteomics and Proteomics 1

Overview	2
1.1 Glycan Building Blocks	3
1.2 N-Glycan Structures	7
1.3 N-Glycan Biosynthesis	8
1.4 Metabolic Engineering of N-Glycans.....	11
2.1 Biological Roles of Glycans on Immune Cells and Antibodies	12
2.2 Changes in Antibody Glycans in Diseases	14
2.3 Sample Preparation for N-glycomic Analysis of Cell Membrane	16
2.3 Sample Preparation for N-glycoproteomic Analysis of Cell Membrane	19
3.1 Fundamentals of Liquid Chromatography (LC) and Mass Spectrometry (MS)	19
3.2 LC-MS/MS Analysis of N-Glycans	24
3.3 LC-MS/MS Analysis of Peptides and Glycopeptides	27
References.....	32

Chapter 2. Investigating The Effects of Glycosylation Inhibitors and Monosaccharides on Cell Surface and Secreted Antibody (IgG1) Glycosylation of CESS Lymphoblast Cell Line Using LC-MS/MS..... 41

Abstract.....	42
1. Introduction	43

2. Materials and Methods.....	46
2.1 Cell Culture and Treatments	46
2.2 Cell Surface Glycomics	46
2.3 Sample Preparation for N-Glycan Analysis.....	47
2.4 N-Glycan Analysis.....	48
2.5 IgG Purification.....	49
2.6 IgG Digestion	49
2.7 Glycoproteomic Analysis With LC-MS/MS	50
2.8 Glycoproteomic Data Analysis	50
2.9 IgG1 Molecular Modeling and Molecular Dynamics	51
2.10 RNA Extraction and 3'-TagSeq RNAseq Analysis	51
3. Results.....	52
3.1 N-Glycomics of the Cell Membrane	53
3.1.1 The Effects of Glycosyl Transferase Inhibitors on the N-glycome of the Cell Membrane....	53
3.1.2 The Effects of Monosaccharide Supplementations on the N-Glycome of the Cell Membrane	57
3.2. Glycoproteomics Analysis of the Secreted IgG1	59
3.2.1 The Effects of Glycosylation Inhibitors on Secreted IgG1 Glycosylation	60
3.2.2 The Effects of Monosaccharide Supplementations on Secreted IgG1 Glycosylation	62
3.3 Modeling IgG1 Fc interaction with IgG1 Fc Receptor	64
3.4 Transcriptomics Analysis of Glycosylation Inhibitor Treated Cells	66
4. Discussion.....	67
5. Conclusion.....	69
References.....	70
Supplementary Figures	75

Chapter 3. Pemphigus-associated desmoglein-specific IgG1 and IgG4 have a dominant agalactosylated glycan modification	82
Abstract.....	83
1. Introduction	84
2. Materials and Methods.....	85
Serum Sample Preparation	86
Indirect Immunofluorescence (IIF).....	86
Desmoglein ELISA.....	86
UPLC-ESI-QqQ-MS Analysis of Pemphigus Serum Samples	87
Enrichment of DSG 3-Specific Antibodies	88
Glycoproteomic Analysis with Orbitrap LC-MS/MS	88
Analysis of Orbitrap LC-MS/MS data.....	89
Statistical analysis	89
Data Sharing.....	90
3. Results	90
3.1 Glycoproteomics of antibodies in pemphigus vulgaris patients using MRM LC-MS/MS	90
3.2 Comparison of glycosylation changes in antibodies of pemphigus vulgaris (PV) and primary biliary cirrhosis (PBC) patients	92
3.3 Glycoproteomics of disease specific autoantibodies in pemphigus vulgaris patients using ELISA and LC-MS/MS.....	92
3.4 Correlation of sialylated Ig glycans with autoantibody titers in pemphigus vulgaris	93
4. Discussion.....	95
5. Conclusion	96
References.....	97
Supplementary Figures	100
Supplementary Figures	105

Chapter 4. Absolute Quantification of Over 600 Bovine Milk Proteins Using a Novel nanoLC-MS/MS Based Bottom-Up Proteomics Approach.....	109
Abstract.....	110
1. Introduction	111
2. Materials and Methods.....	113
Bovine Milk Total Protein Concentration Measurements with BCA	113
Bovine Milk Protein Digestion Protocol	113
Bovine Milk Protein Standards Digestion Protocol	113
Tryptic Digest Clean-up Protocol	114
LC-MS/MS Analysis	114
Proteomics Identification and Quantitation	115
GO Enrichment Analysis	116
3. Results and Discussion	117
Conventional vs Novel Bovine Milk Protein Quantification	117
Validation of Absolute Protein Concentrations.....	122
Absolute Concentrations of Bovine Proteins in Milk	127
Quantified Bovine Milk Proteins Have Potential Health Benefits.....	135
4. Conclusion.....	138
References.....	139
Supplementary Information.....	143

Chapter I

Introduction to Glycans, Immunoglobulins and Liquid Chromatography-Mass Spectrometry Methods for Glycomics, Glycoproteomics and Proteomics

Overview

Glycans are essential biomolecules that play pivotal roles in cell-cell communication, signaling, structural support and immune system modulations. Comprising of various sugar monomers with unique glycosidic linkages, branching, and potential modifications, glycans produce diverse structures. This diversity in structures enables them to perform specialized functions, making them integral components of glycoproteins, glycolipids, and proteoglycans in living organisms.

Aberrant glycosylation is observed in diseases as diverse as cancer and auto-immunity. Deducing glycan structures is especially important for unraveling their roles in these diseases. Additionally, many antibody therapeutics are also glycosylated. Therefore, deciphering the glycan structures associated with these drugs are essential steps for optimizing their function.

Advances in analytical instruments, in particularly in Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS), have facilitated the detection and identification of glycans and glycopeptides. In this research, these instruments were essential tools for detecting glycan changes in tissue or serum due to diseases. LC-MS/MS techniques were also used for monitoring changes in glycosylation of cell membrane after treatment with various nutrients, in particular monosaccharides.

1.1 Glycan Building Blocks

Glycans are composed of monosaccharides that are often covalently bonded to proteins and lipids. Common glycan building blocks include D-glucose (Glc), N-acetyl-D-glucosamine (GlcNAc), D-galactose (Gal), N-acetyl-D-galactosamine (GalNAc), D-mannose (Man), L-fucose (Fuc), N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc) (Figure 1.1).¹ The monosaccharides are converted into their activated forms (UDP-, GDP- or -CMP) in the cytoplasm of cells before being incorporated into glycans by a series of glycosyltransferases.²

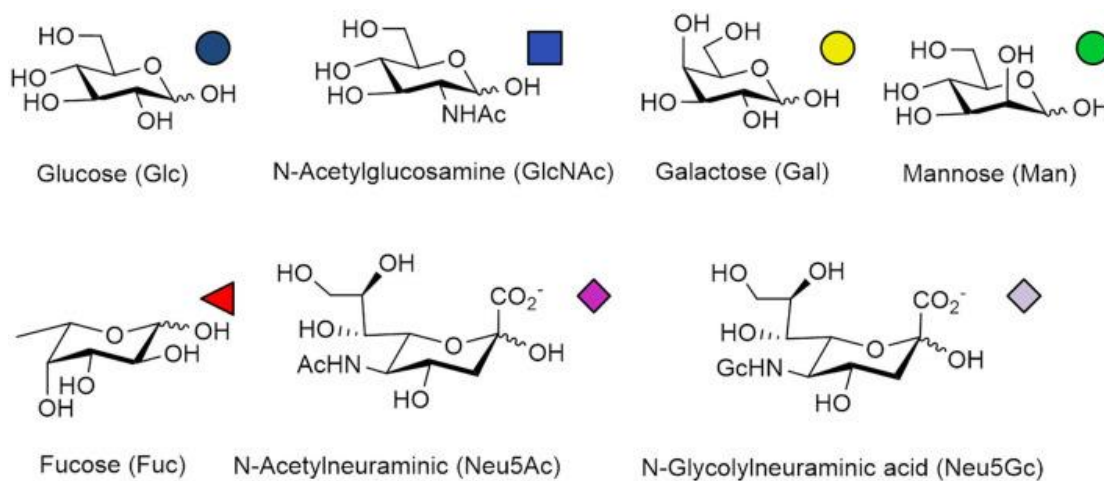


Figure 1.1. Common monosaccharides in N-glycans. Adapted from Ref.¹

Monosaccharide precursors for glycan biosynthesis are derived externally and from intracellular sources.³ Externally derived monosaccharides are brought into cells using two types of transporters, namely energy-independent facilitated diffusion and energy-dependent.³ In humans, the most common glucose transporters (GLUT) are the former, while sodium-dependent glucose transporters (SGLT) in intestinal and kidney epithelial cells are examples of the latter.⁴ In the cell feeding trials below, these transporters play a significant role for creating the building

blocks necessary for glycan production. Thus, externally derived glucose gets converted to GlcNAc in a series of enzymatic processes.⁴

One of the major intracellular sources for building glycans is the salvage of monosaccharides from degraded glycans in cell lysosomes. Glycan degradation takes place in the lysosome, and it yields the monosaccharides that upon activation can be directly incorporated in structures.⁴ Glycan degradation through the salvage pathway and its contribution to glycosylation are significant. For example, it has been shown that nearly 80% of the GlcNAc from glycoproteins degraded in liver lysosomes is converted into UDP-GlcNAc, and at least 30% is used again to synthesize secreted glycoproteins.⁴ After degradation, monosaccharides released by the lysosome into the cytoplasm are generally activated and reused as shown in Figure 1.2.⁵

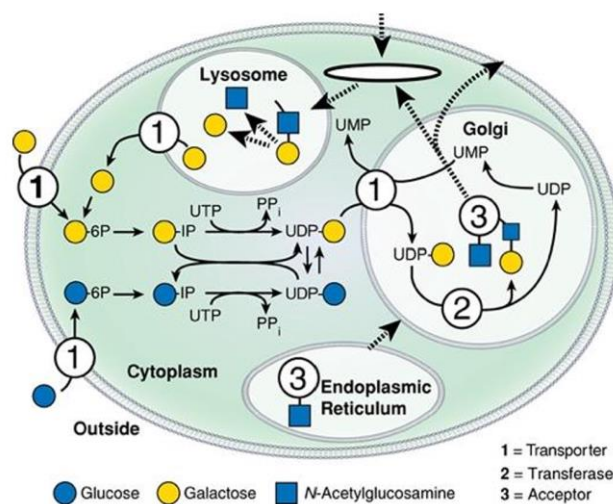


Figure 1.2. Glycan degradation produces glycan building blocks which turn into glycans. Adapted from Ref.⁵

Another major intracellular source of glycan building blocks is the interconversion pathway. Interconverting monosaccharides from one to another produces specific monosaccharides as building blocks. The biosynthesis of these conversions are schematically shown in Figure 1.3.⁴ Glucose is the most common monosaccharide for use in the interconversion pathways as it can be converted into all glycan building blocks.⁴ Glucose can even be converted to monosaccharides such as fucose and NeuAc, which are unique residues often associated with diseases.^{6,7} Fucose and NeuAc are important residues and can be generated through both monosaccharide interconversion and salvage pathways.⁴

Understanding the synthesis of the monosaccharide building blocks and identifying the regulatory points in the glycosylation process are crucial for enhancing the metabolic engineering and inducing changes in glycosylation of membrane or secreted proteins. However, predicting the effects of providing monosaccharides to cells *in vitro* on cellular glycosylation can be challenging due to the complex interconversion pathways of monosaccharides. In previous studies, we demonstrated the diverse effects of different monosaccharides on N-glycans of epithelial cells.⁸

Mutations affecting regulatory points in glycan biosynthesis have significant consequences, such as the inability to synthesize and integrate specific monosaccharides into glycans. For instance, human glycans lack N-Glycolylneuraminic acid (Neu5Gc), a common mammalian glycan component, due to a mutation in the CMAH gene.⁹ CMAH gene encodes the hydroxylase enzyme responsible for converting CMP-Neu5Ac to CMP-Neu5Gc.¹⁰ Neu5Gc is typically found in mammals lacking the mutation in CMAH gene. Consequently, many pathogens target Neu5Gc for invasion.⁹

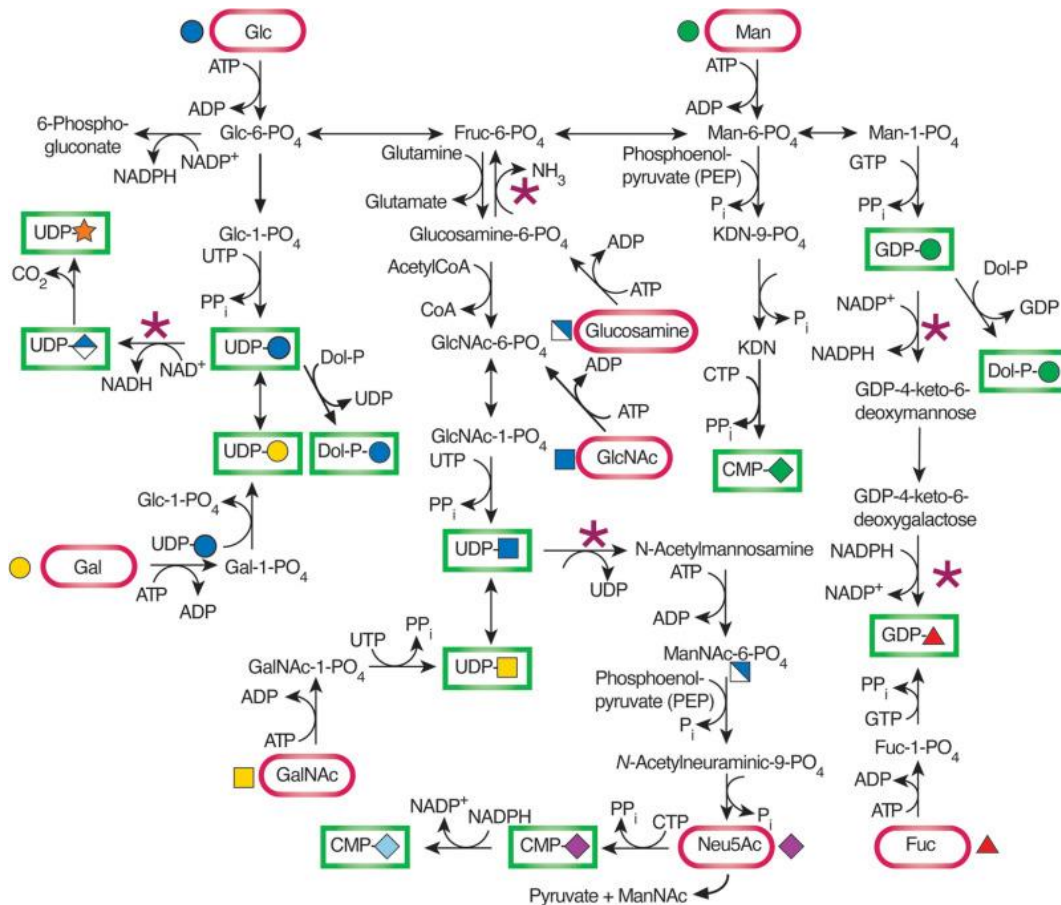


Figure 1.3. The biosynthetic pathways for interconversions of the monosaccharides observed in eukaryotes. Adapted from Ref.⁴

Glycans are often covalently bonded to proteins and lipids. Proteins that contain glycan moieties are known as glycoproteins, while glycan containing lipids are referred to as glycolipids.^{11,12} Glycoproteins can be further classified into two main groups based on the glycan attachment sites into two main groups: N-glycans, which are attached to asparagine amino acids, and O-glycans, which are attached to serine or threonine amino acids.^{13,14}

1.2 N-Glycan Structures

N-Glycans are covalently linked to proteins through asparagine residues (N) within the consensus sequence Asn-X-Ser/Thr, where X can be any amino acid except proline.¹³ Glycans exhibit significant structural diversity which are influenced by the specific glycoprotein and the expressing organism.¹⁵ A commonality of all N-glycans is their core structure, which consists of two N-acetylglucosamine (GlcNAc) residues, and three mannose (Man) residues.¹³ The variability in N-glycan structures primarily arises from the processing and modifications to its core structure. These modifications can include branching, fucosylation, sialylation, and galactosylation.¹³ N-Glycans can be further classified into three types based on their structures (Figure 1.4).¹³ High mannose, also known as oligomannose, are N-glycans that contain mannose residues in continuation of the N-glycan core. In complex N-glycans, both the α 3- and α 6-linked mannose residues are substituted with GlcNAc moieties, whereas hybrid N-glycans have both GlcNAc substituted and unsubstituted mannose residues (Figure 1.4).¹³

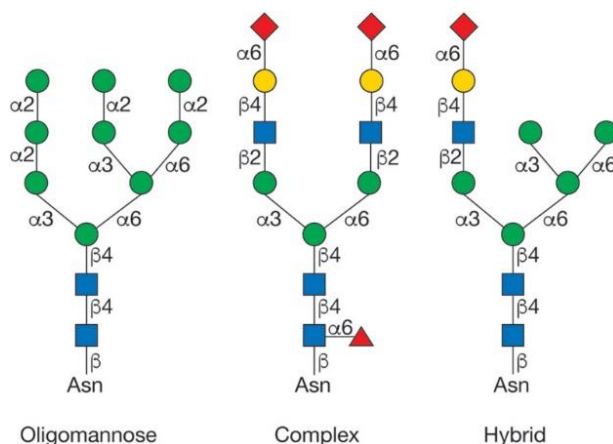


Figure 1.4. Three major types of N-glycans. Adapted from Ref.¹³

Furthermore, N-glycans can be classified based on containing specific residues (Figure 1.5).¹⁶ The glycans that have undergone high mannose trimming and GlcNAc addition are considered undecorated. The fucosylated glycans contain fucose and the sialylated glycans contain sialic acid. Lastly, the glycans with both sialic acid and fucose are known as sialofucosylated N-glycans.

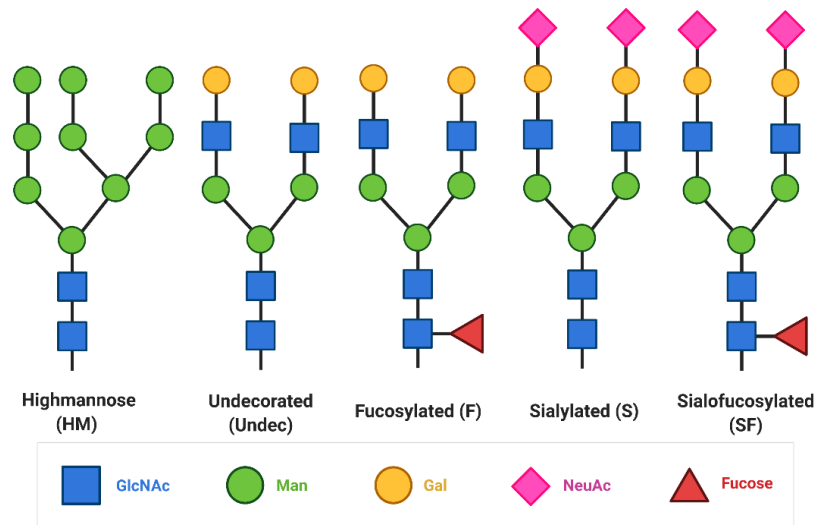


Figure 1.5. N-Glycan sub-type classifications based on their decorations.

1.3 N-Glycan Biosynthesis

The biosynthesis of N-glycans is non-template driven, unlike protein biosynthesis.¹⁷ In eukaryotes, glycosylation occurs inside the endoplasmic reticulum (ER) and Golgi apparatus.¹³ Glycosyltransferases carry the activated forms of monosaccharides and transfer the saccharide moieties to the proteins in this process.⁵

N-Glycan biosynthesis (Figure 1.6) is initiated by the dolichol pathway. GlcNAc-1-phosphotransferase, DPAGT1, transfers GlcNAc from uridine diphosphate N-acetylglucosamine

(UDP-GlcNAc) to dolichol phosphate, which leads to formation of N-acetyl- α -D glucosaminyl-diphosphodolichol.¹³ Thereafter, an additional GlcNAc, nine mannose residues, and three glucose residues are added to dolichol to form the N-glycans precursor, Glc3Man9GlcNAc2-P-P-Dol.¹³ This glycan precursor is then placed on an asparagine residue of a peptide. The three glucose residues of the precursor are then trimmed off and the resulting high-mannose glycoprotein structure gets moved to Golgi to be further processed by high-mannose glycan trimming enzymes.¹³

Figure 1.6. Biosynthesis of Glc3Man9GlcNAc2-P-P-Dol, which is the precursor for all N-glycans.

High-mannose glycans are the foundation of all types of N-glycans and are the least processed. After mannose residues of high mannose glycans are trimmed in the Golgi apparatus, GlcNAc, galactose, and other residues such as fucose or NeuAc are added on the glycan to create the final products (Figure 1.7).¹³ The step-by-step systematic biosynthesis of N-glycan enables the use of metabolic engineering tools to induce changes in N-glycans. These changes are produced by inhibiting certain steps of the process or providing additional glycan building blocks to the cells.

1.4 Metabolic Engineering of N-Glycans

Glycosylation enzymes, or glycoenzymes, are responsible for trimming the mannose residues on high-mannose structures and adding other monosaccharide building blocks to the nascent structure. Specific glycoenzymes can be inhibited to induce selective changes in the final glycan products (figure 1.8). Kifunensine is an alkaloid with a high inhibitory effect on class I-mannosidases, such as MAN1A1, MAN1A2, MAN1B1, and MAN1C1.¹⁸ These mannosidases trim mannose residues of Man-9, turning it into Man-8, Man-7, Man-6, and subsequently Man-5.¹⁹ Man-8, Man-7, Man-6, and Man-5 have 8, 7, 6, and 5 mannose residues respectively. High mannose glycans are generally not fucosylated or sialylated, as the respective enzymes act later in the process. Therefore, inhibiting mannosidases with kifunensine leads to production of mostly high-mannose glycans with no decorations.²⁰

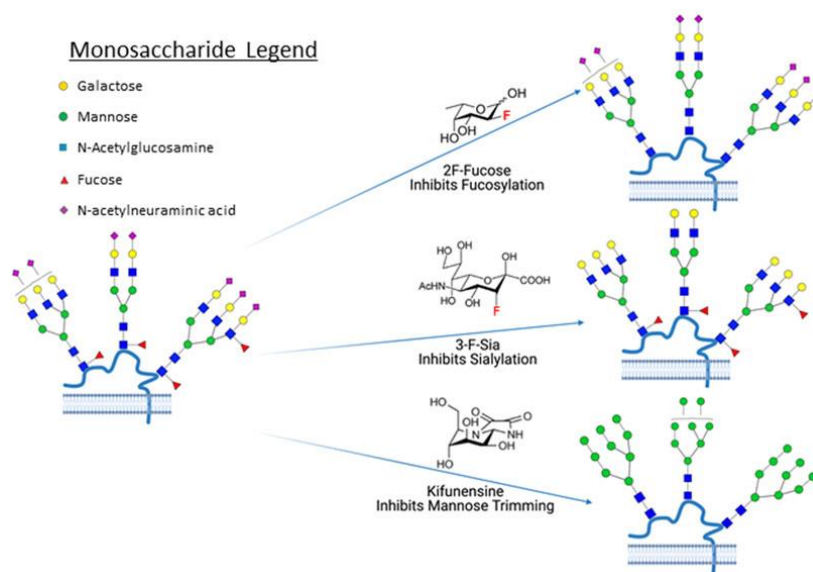


Figure 1.8. Induced changes in cell surface N-glycan using glycoenzyme inhibitors. Adapted from Ref.²⁰

Like mannosidase enzymes, decorating glycoenzymes may be inhibited to induce glycan changes as well. Fucosylation may be inhibited by a fluorinated analogue of fucose, such as 2-F-Fucose.²⁰ 2-F-fucose works by interfering with formation of UDP-fucose in the cells.²¹ While, a fluorinated analogue of sialic acid, such as 3-F-Sia, disables the sialylation of N-glycans by binding to sialyltransferases after it is converted to CMP-P-(3-F-Sia).²⁰ Glycosylation inhibitors are useful because they enable selective changes of cell membrane glycans. After these selective glycan changes, biological functions of cells and glycoproteins, such as binding, may be assessed to further understand the roles of specific glycans on these functions. Using these techniques, we demonstrated sialylated N-glycans affect the binding of Spike RBD protein to the cell surface of epithelial cells *in vitro*.²⁰

2.1 Biological Roles of Glycans on Immune Cells and Antibodies

Glycans are especially important in functionality of immune cells, such as B cells.²² B cells are a type of white blood cells that are involved in both innate and adaptive immunity.²³ Cell surface glycans of B cells enable their interactions with T cells and antigen-presenting cells.^{24,23} These interactions are essential for coordination of immune responses. Additionally, other immune cell functions, such as migration to infected sites to cause inflammation, are dependent on glycan-protein interactions.²⁵ Selectin-mediated adhesion of leukocytes to endothelial cells is driven by glycan-protein interactions that facilitate lymphoblast trafficking.²⁶

Immunoglobulins, or antibodies, are glycoproteins produced by B cells. All antibodies are heavily glycosylated with N-glycans and even O-glycans (Figure 1.9).²⁷ N-Glycans on antibodies influence their stability, effector functions, and interactions with other immune cells and

antigens.^{28,29,23} These glycans modulate antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC), which are the mechanisms used for clearing pathogens.³⁰

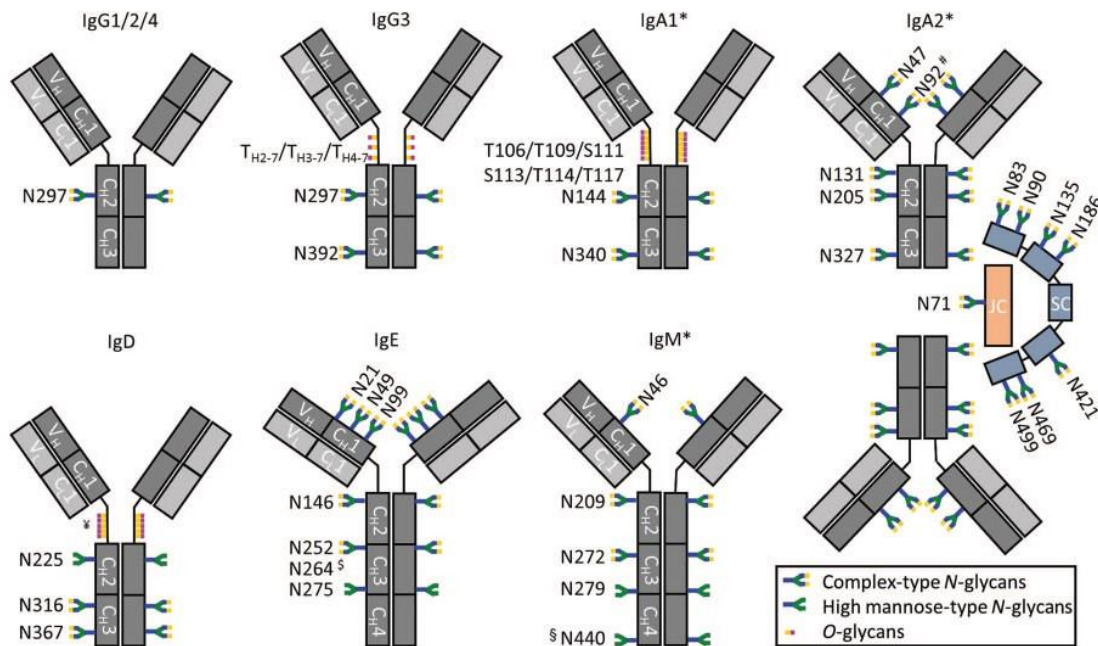


Figure 1.9. Glycans on antibodies (IgG, IgA, IgD, IgE and IgM). Adapted from Ref.²⁷

CH84.4 is a conserved glycosylation site on IgG and specifically corresponds to Asn-297 of IgG1. The glycans at this site affect IgG1 binding to Fc receptors, hence IgG1 is classified as pro- or anti- inflammatory depending on the types of the glycans at this site (Figure 1.10).²⁹ Specific glycans at this site, such as the ones with core fucose, galactose, and sialic acid have reduced binding to Fc receptors.²⁹

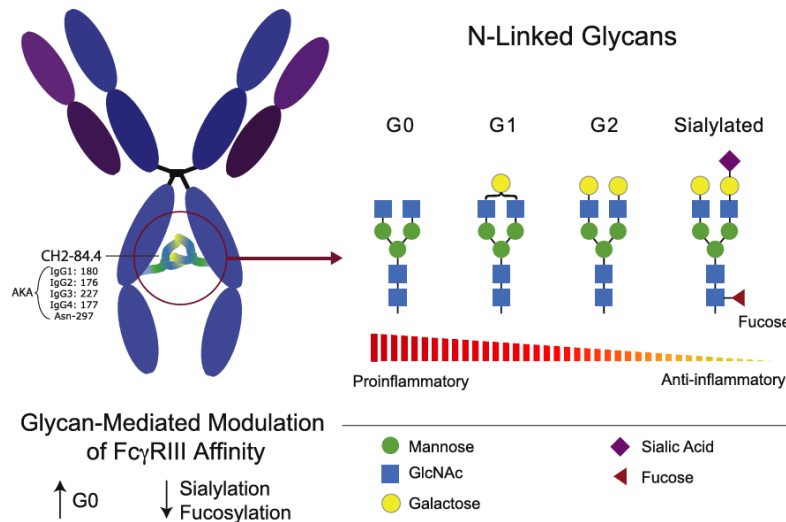


Figure 1.10. Effects of different N-linked glycans on IgG inflammatory state. Adapted from Ref.²⁹

2.2 Changes in Antibody Glycans in Diseases

Glycosylation changes of antibodies are observed in various diseases. We have demonstrated these changes in gastric and lung cancers, and in Alzheimer's disease.^{31,32,33} Furthermore, major shifts in antibody glycans are observed with auto-immune diseases.²⁹ The details of these disease-specific changes are presented in Table 1.1.²⁹ Glycan changes can be used as biomarkers for early detection of these diseases. Additionally, these changes often contribute to disease progression.²³ The altered glycan theory of autoimmunity hypothesizes that each auto immune disease has a unique N-glycan profile due to the importance of N-glycans in antibody effector functions.²⁹ To better understand the physiopathology of auto-immune diseases, it is crucial to characterize the glycosylation profiles of antibodies in all auto-immune diseases, especially the ones with disease-specific antibodies.

Table 1.1. Variations in IgG glycans. Adapted from Ref.²⁹

Disease	Glycan Modification
Rheumatoid Arthritis	Decreased N-linked galactosylation and sialylation of IgG ³⁴ Decreased N-linked galactosylation and terminal sialylation of ACPA-IgG1 ³⁵ Increased fucosylation of ACPA-IgG1 ³⁵
IgA Nephropathy	Decreased terminal O-linked galactosylation on the hinge region of IgA1 heavy chains ^{36–39}
Gastric Cancer	Decreased galactosylation and sialylation of IgG ³¹
Henoch-Schonlein Purpura	Decreased galactosylation and sialylation of IgA1 ³⁶
HIV	Decreased galactosylation, sialylation, and fucosylation in HIV-specific antibodies associated with enhanced Fc-mediated reduction of viral replication and enhanced Fc receptor binding ⁴⁰
Influenza and Tetanus Vaccination	Increased galactosylation and sialylation of anti-vaccine IgG Increased number of sialic acid per glucose during vaccination Decreased IgG1 bisecting N-acetylglucosamine ⁴¹
Lambert-Eaton Syndrome	Decreased galactosylation of IgG1 and IgG2 Increased levels of bisecting N-acetylglucosamine on IgG1 and IgG2 in patients less than 50 years old ⁴²
ANCA Vasculitis	Decreased galactosylation of IgG ⁴³ Decreased sialylation of IgG Fab'2 ⁴⁴
Myasthenia Gravis	Decreased galactosylation of IgG2 ⁴²
Ovarian Cancer	Increased agalactosylated biantennary glycans on IgG heavy chains ⁴⁵
PR3-ANCA Vasculitis	Decreased N-glycan 2,6-linked sialylation ⁴⁶ Decreased galactosylation of IgG ⁴³
Sjogren's Syndrome	Decreased sialylation of IgG ⁴⁷

2.3 Sample Preparation for N-Glycomic Analysis of Cell Membrane

Glycosylation of cell membranes and glycoproteins can be characterized through glycomics and glycoproteomics.¹⁶ In glycomics, N-glycans are released using the enzyme PNGase F which cleaves off the N-glycans from asparagine residues (Figure 1.11). PNGase F works on all types of N-glycans except the ones with alpha 1-3 core fucose, which are generally found in plants.⁴⁸ Once released, glycans may be cleaned or de-salted using solid-phase extraction (SPE) with porous graphitic carbon (PGC) as the stationary phase. After clean-up, the samples are evaporated to dryness and reconstituted in water to a final concentrations before direct injection into an LC-MS/MS for analysis.¹⁶

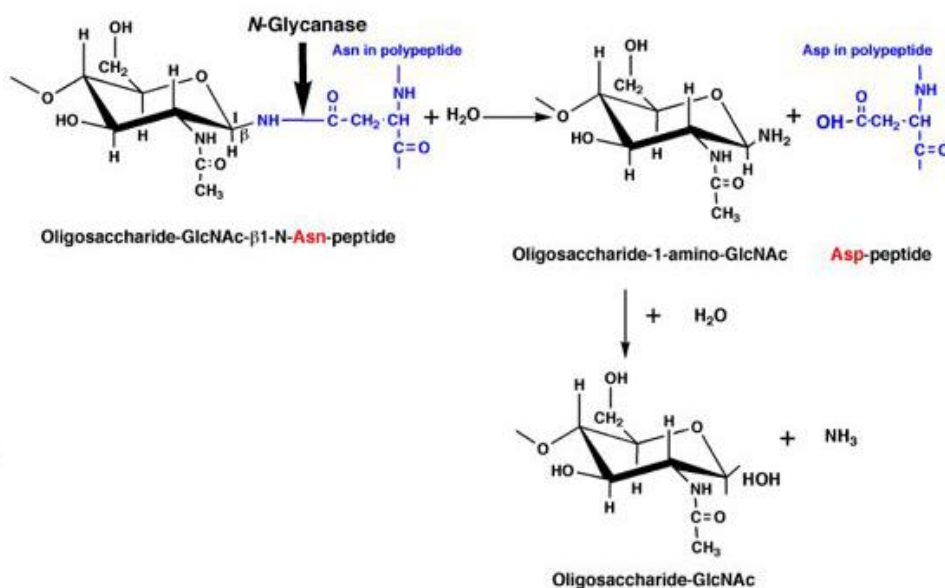


Figure 1.11. PNGase F (N-Glycanase) cleaves off the N-glycans from Asn residues. Adapted from Ref.⁴⁹

For analysis of cell membrane N-glycans, the cells are initially harvested, and the cell membrane is enriched. These steps are shown schematically in Figure 1.12. Adherent cells are

generally scraped from petri dishes in DPBS for harvesting,¹⁶ while suspension cells are centrifuged. After harvesting the cells, they are lysed by sonication. The cell debris is then removed through centrifugation. The cell membrane is separated in a second phase with ultracentrifugation by collecting them at the bottom of the vial away from the cell cytoplasm.¹⁶ The membrane proteins need to be denatured prior to N-glycan release. Therefore, the membrane glycoproteins are treated with a reducing agent, such as DTT, and heated to break down the disulfide bonds and denature the glycoproteins. Afterwards, PNGase F enzyme is added with overnight incubation in a 37°C water bath to cleave off the N-glycans. N-glycans are cleaned or de-salted prior to analysis using LC-MS/MS, as impurities and salts may interfere with the analysis. N-Glycans are often cleaned using PGC SPE prior to LC-MS/MS analysis.¹⁶

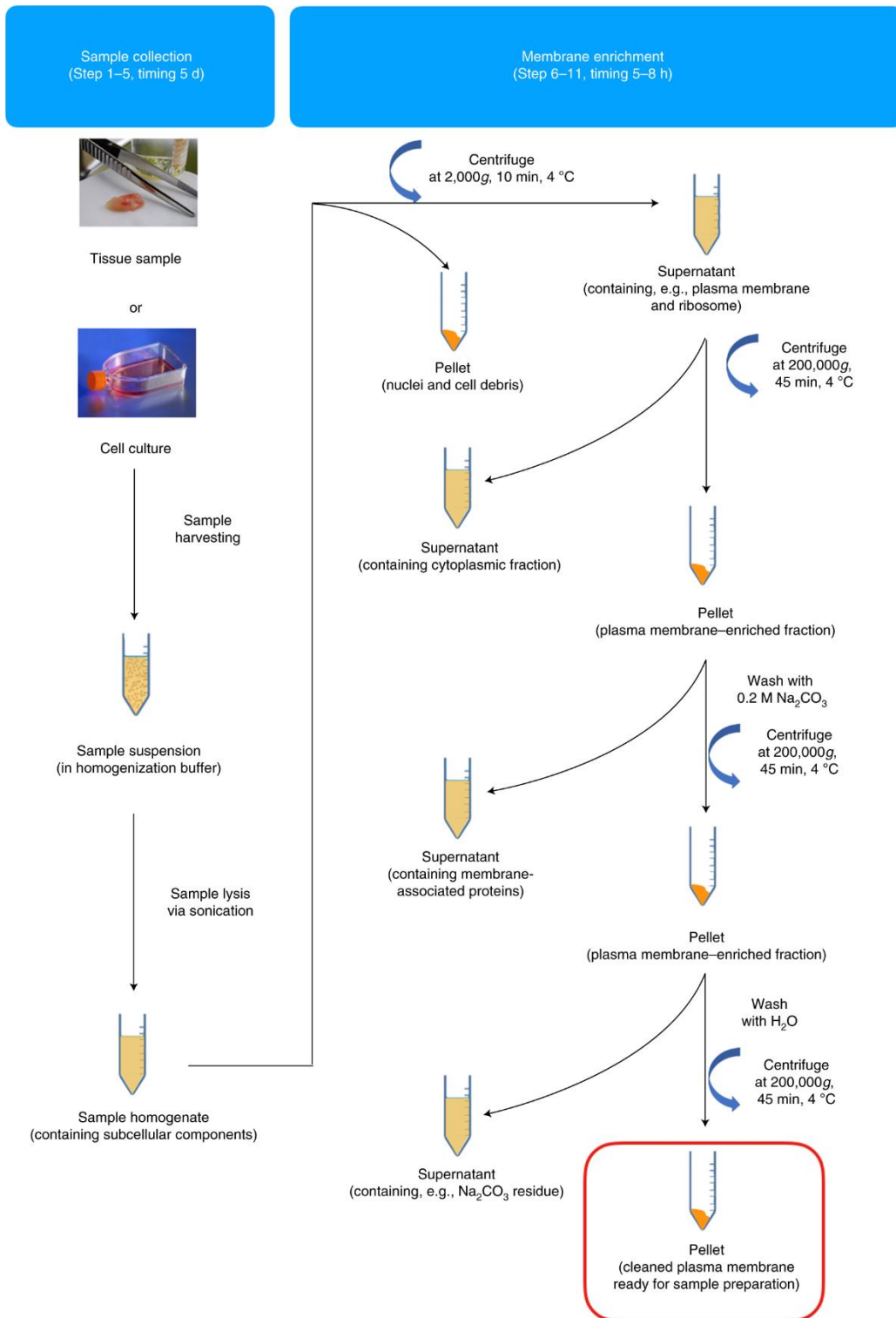


Figure 1.12. Schematic workflow for cell membrane extraction from cells or tissues. Adapted from Ref.¹⁶

2.4 Sample Preparation for N-Glycoproteomic Analysis of Cell Membrane

Glycoproteomic analysis yields the identity of glycoproteins along with the glycosylation sites and glycan compositions. To perform this analysis on the cell membrane glycoproteins, the cells are harvested and lysed as with glycomics analysis. After capturing the cell membrane through ultracentrifugation, the glycoproteins undergo denaturation using DTT reduction to breakdown the disulfide bonds and IAA alkylation to cap the cysteine residues.⁵⁰ Then, a protease is used to break down the glycoproteins into smaller peptides and glycopeptides.¹⁶ Trypsin is a commonly used protease and cleaves the peptide backbones at C-terminal of lysine and arginine residues.⁵¹ Glu-C protease is often paired with trypsin for glycoproteomics because it cleaves C-terminals of aspartic and glutamic acid residues⁵² and produces smaller glycopeptides and peptides. After digestion, peptides make up most of the mixture, therefore, glycopeptide enrichment is crucial for improved identifications. Hydrophilic interaction chromatography (HILIC) SPE is capable of enriching glycopeptides, as it retains glycopeptides more than peptides. Peptides may be washed away using HILIC SPE cartridges thereby reducing their abundance to less than 10%.¹⁶ After this enrichment step, glycopeptides are analyzed using LC-MS/MS.¹⁶

3.1 Fundamentals of Liquid Chromatography (LC) and Mass Spectrometry (MS)

A mass spectrometer is a type of detector on chromatographic instruments that determines the mass of compounds. The measured masses and fragmentation data obtained by a mass spectrometer are often compared to mass libraries for identification of compounds. Peak intensities in a MS output, known as a mass spectrum, are also used for quantification purposes.

These attributes make mass spectrometry a powerful tool for qualitative and quantitative analysis of bio-analytes such as glycans and peptides.¹⁶

Prior to MS analysis, compounds are typically separated using two major chromatography techniques: liquid chromatography (LC) and gas chromatography (GC)⁵³. Gas chromatography offers fast separation with high efficiency,⁵⁴ but it has limited use because most biologically derived compounds have high vaporization temperatures. Alternatively, LC is more versatile and is compatible with almost all water soluble compounds especially biomolecules.⁵⁵

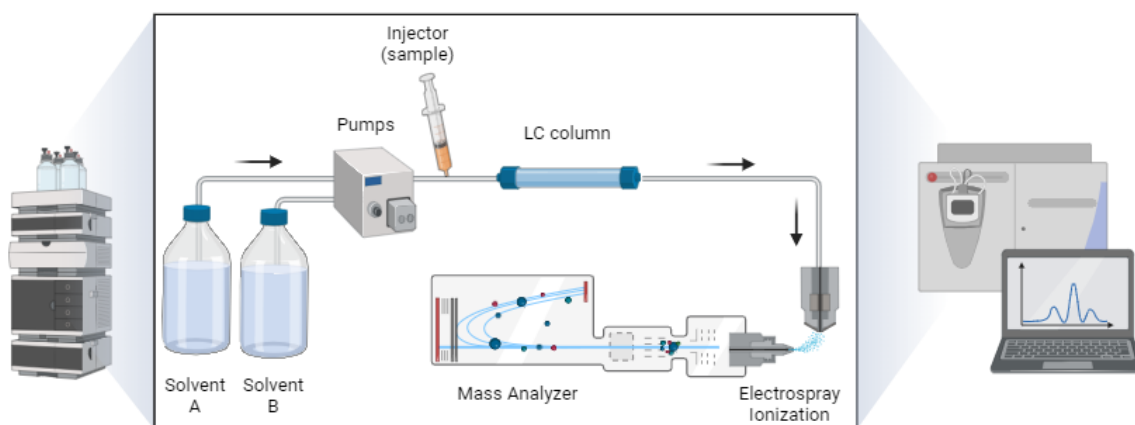


Figure 1.13. The general scheme of a LC-MS.

LC uses a stationary phase and a mobile phase, which interact differently with the compounds depending on its polarity. Some compounds have stronger affinities for the stationary phase and will be retained longer in comparison with the compounds that have lower affinities for the stationary phase. The mobile phase simply carries the compounds through the column and may have different compositions at different times of analysis. The variations in mobile phase composition throughout the analysis is known as the gradient.⁵⁶

LC is often operated in the reversed-phase mode, where the stationary phase is made of a hydrophobic or nonpolar compound, such as porous silica particles linked to alkyl chains (e.g. C18).⁵⁷ C18 stationary phase is commonly used for separation of glycopeptides and peptides due to relatively high hydrophobicity of the peptide moieties.⁵⁸ In the gradient used for reversed-phase LC, the mobile initially phase contains a higher percentage of a polar solvent such as water, but as the separation takes place, a less polar solvent such as acetonitrile makes up a larger component of the mobile phase in order to elute more hydrophobic compounds that were retained by the column.⁵⁷ As shown schematically in Figure 1.13, the compounds of interest are carried to the column by the mobile phase, and as they exit the column they enter the ionization source for MS analysis.

The major components of a mass spectrometer are the ion source and the mass analyzer detector. Ionization is an important step in MS analysis because compounds need to be charged to be detected by MS. For this research, electrospray ionization (ESI) was the main method for ionization. In ESI, a high voltage is applied to the sample at the nebulizer to induce formation of charged macro particles (Figure 1.14). This takes place in the presence of gases at elevated temperatures, nitrogen at 300 °C, which helps with solvent evaporation. As the charged particles decrease in size, they undergo coulomb explosions to decrease the charge repulsion in the droplets.⁵⁹ Eventually, the particles break down into very small particles and eventually charged analytes, which are then detected by MS.

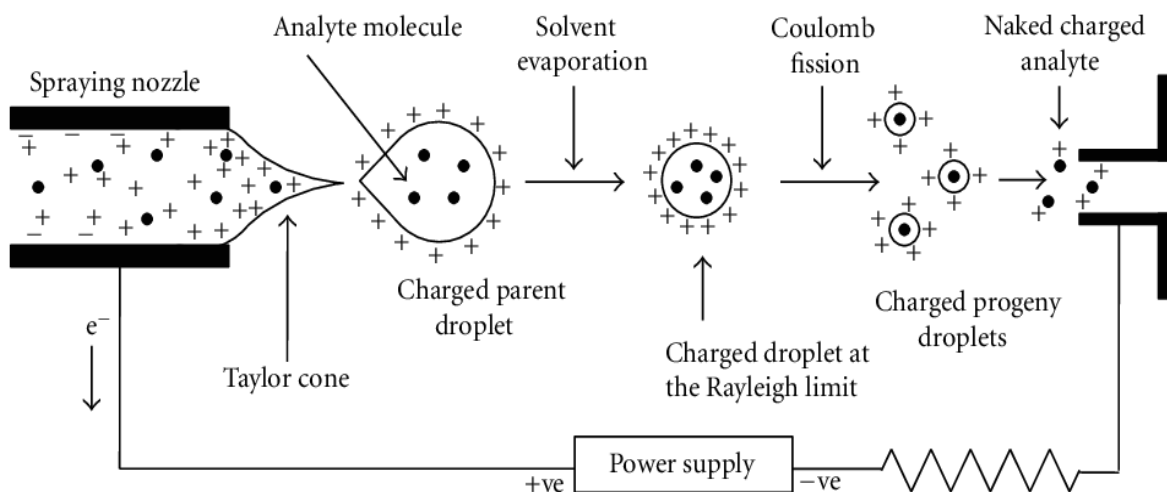


Figure 1.14. Soft ionization by electro spray ionization. Adapted from Ref.⁵⁹

Mass analyzers apply either magnetic or electric fields or both to manipulate ions through a specific path in the device. For the work described here, two types of mass analyzers were used. Time-of-flight (TOF) is a mass analyzer that measures the mass to charge ratio of the ion based on the travel time.⁶⁰ Orbitrap mass analyzers are newer and more complicated but have higher resolution and sensitivity compared to the previous generations of MS. Due to their complexities, these instruments are described in greater detail here.

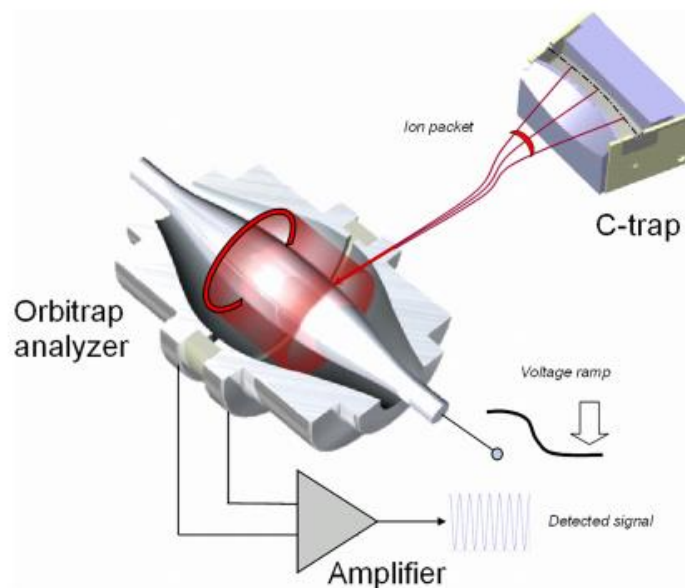


Figure 1.15. Injection of ion packets into orbitrap analyzer. Adapted from Ref.⁶¹

In orbitraps, ions are forced to cycle around a central electrode while simultaneously oscillating along a horizontal axis.⁶² Central and outer electrodes of orbitrap are shaped to induce both the horizontal and oscillating motions of ions. The two electrodes are controlled by independent voltages. By applying an electric field between the inner and outer electrodes, the ions are attracted to the inner electrode and those that have sufficient tangential velocity are forced to cycle around the inner electrode.⁶¹ The outer electrode of orbitrap is divided into ion excitation and detection components with the area between the two electrodes, known as the measurement chamber, being under ultrahigh vacuum. Coherent movement of ions is required to receive high resolution signal. Therefore, the ions are injected in short bursts as ion packets (Figure 1.15).⁶¹ To introduce the ions into orbital trapping efficiently, the electric field of the injection port is continuously modified by a field compressor in the outer electrode.⁶¹

Mass detection in orbitraps is carried out by measuring the coherent oscillations in axial direction producing an image current detection.⁶¹ A Fourier Transform (FT) is applied to the time spectrum converting it into a frequency domain spectrum.⁶¹ The frequency domain spectrum is then converted to a mass spectrum using known resonances of standard ions. Orbitrap has tremendous sensitivity and mass resolving power compared to other MS with mass accuracies within the order of parts per million (ppm).⁶¹

3.2 LC-MS/MS Analysis of N-Glycans

Advances in high resolution mass spectrometry combined with advances in chromatography have modernized the field of bioanalytical chemistry by facilitating identification of proteins, glycoproteins and glycans. Liquid chromatography instruments have been improved to handle much higher pressures allowing the use of much smaller particles in analytical columns leading to improved efficiency and better separation resolution of compounds.⁶³ Additionally, utilization of nanoscale chromatography paired with nanoscale electrospray sources have significantly improved the ionization efficiencies of compounds and has made LC-MS a highly sensitive technique. The innovations in MS instrumentations, such as the development of Orbitrap mass spectrometry has enabled scientists to identify thousands of peptides in an analyte of interest within a few hours with enhanced sensitivity and mass resolution.

Nanoscale liquid chromatography (nanoLC) instruments coupled to high resolution MS (HRMS), such as time-of-flight (TOF) mass spectrometers are incredibly useful in characterizing glycans. Due to the large diversity in the number and types of released glycans from the cell membrane, enhanced separation of these compounds prior to mass spectrometry analysis is of

immense importance. Porous graphitic carbon (PGC) is a unique stationary phase used for separation of oligosaccharides, such as glycans.⁶⁴ The mechanism of oligosaccharide retention by PGC is not fully understood, but it has been hypothesized that dipole-induced CH– π interactions of PGC and saccharides are responsible for their retention.⁶⁵ Oligosaccharides are eluted from PGC with acetonitrile as the mobile phase, due to the interruption of the PGC-saccharide interactions.

There are hundreds of possible N-glycan structures that decorate the cell membrane in cells such as Caco-2 (Figure 1.16, a) making high mass accuracy essential for confident identification of these compounds. TOF instruments can measure m/z with less than 20 ppm mass error, making them an ideal choice for analysis of glycans. Coupling a nanoESI source with these instruments ensures all compounds are ionized properly and improves sensitivity. Furthermore, Q-TOF instruments can produce fragmentation spectra for detected ions through collision-induced dissociation (CID). N-Glycan structures are validated using their fragmentation spectra, as they commonly produce 204 and 366 m/z fragments corresponding to their core GlcNAc and GlcNAc-mannose residues, respectively. N-Glycan fragmentation spectra are also used to validate the presence of fucose and sialic acid residues. A mass shift of 146 m/z is observed when an N-glycan loses a fucose residue through CID (Figure 1.16, b).¹⁶ Additionally, sialylated N-glycans often lose sialic acid residues through CID, as apparent by the presence of 292 m/z ions in the fragmentation spectra (Figure 1.16, b).¹⁶

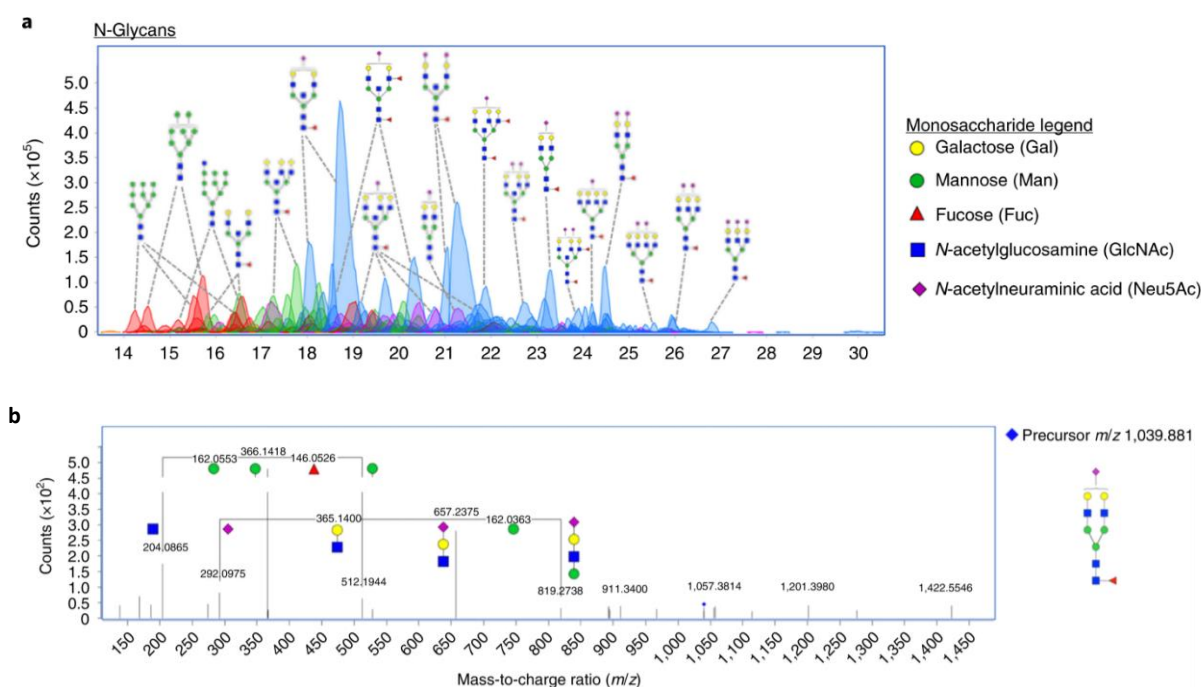


Figure 1.16. a) N-Glycan profile of Caco-2 cell membrane obtained by nanoLC-QTOF MS/MS using a PGC chip yielding identification of more than 500 N-glycans. b) Structural validation of an N-glycan through fragmentation data. Adapted from Ref.¹⁶

N-Glycans are also quantified using LC-MS/MS. Peak areas of identified N-glycans are used to calculate their relative abundances.¹⁶ Relative abundances show the abundance of a compound compared to the abundances of all identified compounds in the sample. Pure N-glycan standards are not commonly commercially available, making it not feasible to obtain absolute quantification (concentrations) of these compounds.

3.3 LC-MS/MS Analysis of Peptides and Glycopeptides

As with the identification of glycans, identification of proteins and glycoproteins have vastly improved with advances in analytical instrumentations. Orbitrap is currently the gold standard for MS in proteomics.⁶⁶ Generally, bottom-up proteomics is the chosen method for identifying proteins in biological samples. In bottom-up proteomics, proteins are digested using proteases, which break down the protein into peptides.⁶⁷ Trypsin is a commonly used protease in proteomics as it cleaves proteins at C-terminus of lysine (K) and arginine (R) residues⁶⁸ resulting in large peptides that are ideal for protein identification. Protein identification is performed by comparing measured masses using LC-MS/MS with the peptide masses of proteins, in addition to validation of peptide sequences using the fragmentation spectra. Therefore, this analysis requires prior knowledge of the protein sequences. Such information may be obtained from protein libraries (e.g. Uniprot) that contain protein sequences based on human, bovine, etc. genomes. A designed workflow for proteomics analysis of bovine milk is presented schematically in Figure 1.17.

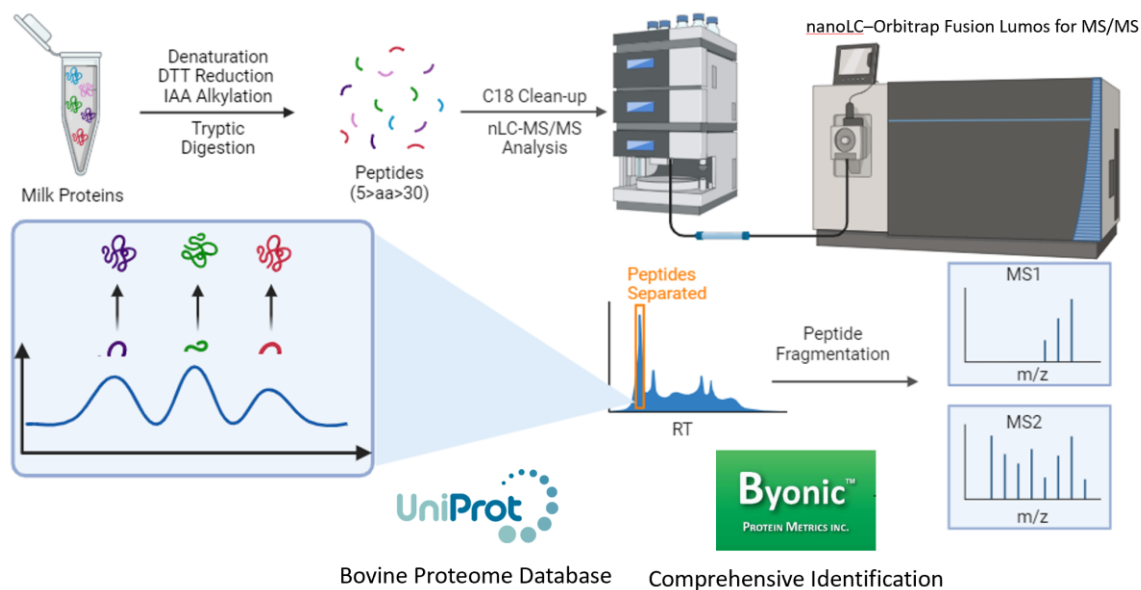


Figure 1.17. Workflow for bottom-up proteomics using nLC-MS/MS.

Even with high mass accuracy of Orbitrap MS (<5 ppm), there are still possibilities of misidentification of peptides because two peptides may have the same amino acid compositions but different sequences. Therefore, peptide identification requires validation with supporting tandem MS (MS/MS) data. Under positive ionization, peptides produce b and y fragments by high-energy collision induced dissociation (HCD),⁶⁹ and these fragment masses can be used to validate the peptide sequences (Figure 1.18).

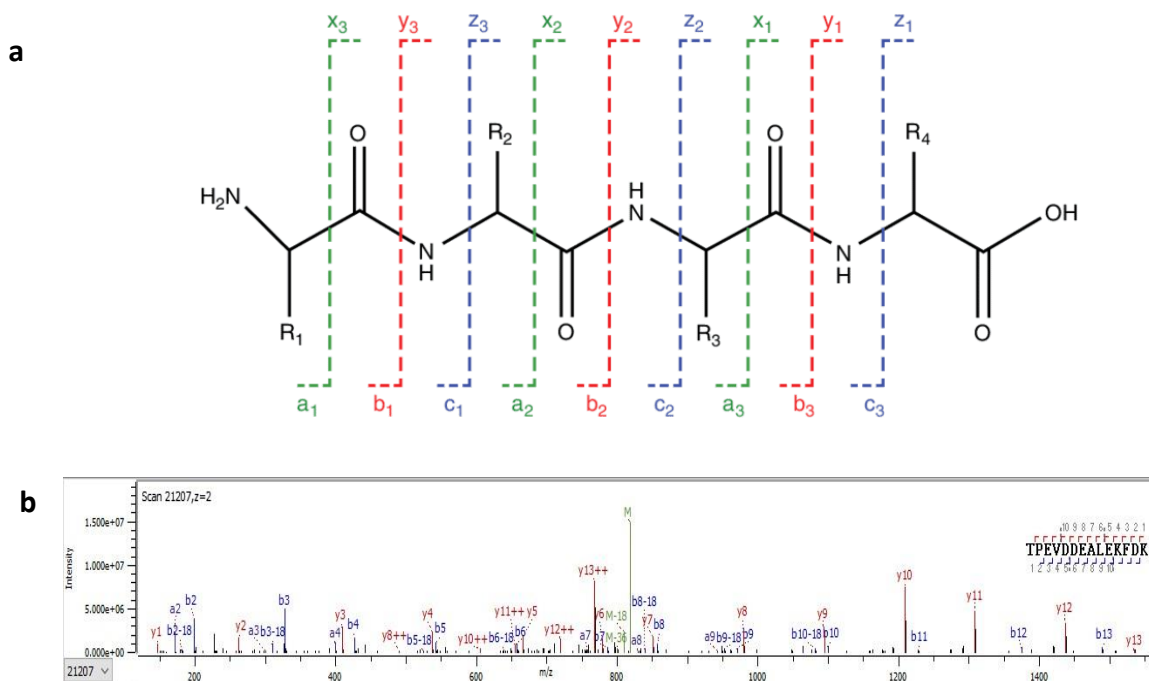


Figure 1.18. (a) Nomenclature for common peptide fragments. Adapted from Ref.⁷⁰

(b) Fragmentation data for 818.3911 m/z precursor mass corresponding to peptide (TPEVDDEALEKFDK) from beta-lactoglobulin protein in bovine milk.

Quantitation of proteins in proteomics has also improved with advances in LC-MS/MS. Methods for absolute quantifications often use protein standards and isotopically labeled peptides as references.⁷¹ These methods are costly and limited because protein standards are not available for most but the most abundant proteins. Additionally, commercial protein standards often contain unique PTMs, which makes the data analysis more challenging. Another approach for quantification of proteins in proteomics uses isotope labeled peptides as references. Peptides have complex isotopic distributions, therefore distinguishing isotope labeled peptides from native

ones is often challenging. Additionally, most commercial software used for proteomics data analysis have not been able to fully address this issue.

Glycopeptides are similarly identified using LC-MS/MS. Having optimal resolution and mass accuracy is of great importance due to the tremendous heterogeneity of glycopeptides. Glycopeptides contain both an amino acid sequence and a carbohydrate moiety, leading to even greater complexity in analysis compared to either peptides or glycans alone. With advanced nanoLC – Orbitrap, these compounds are identified based on both high mass accuracy (<5 ppm) and full fragmentation. An example of the fragmentation spectrum of a glycopeptide is presented in Figure 1.19.⁷² Identification of both the peptide and glycan fragments validate the glycopeptide structures. Given all glycan fragments are not generally observed in this analysis, compositions of glycans are often reported as opposed to their exact structures. With the prior knowledge of glycan biosynthesis, glycan compositions are used to propose putative structures for identified glycans on glycopeptides.

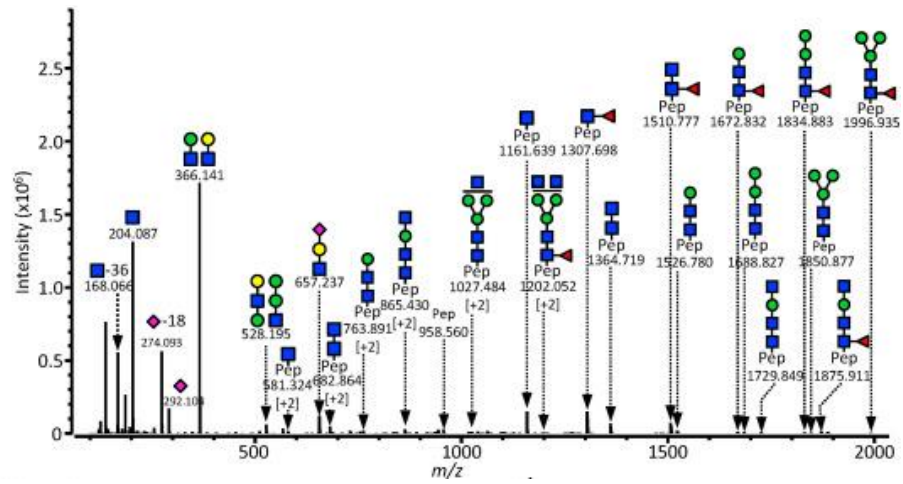


Figure 1.19. Fragmentation data for a glycopeptide with HexNAc5Hex5Fuc1NeuAc1 glycan composition and TPLTANITK peptide sequence (1074.4640 m/z , $z=3$) of sIgA2 in human milk.

Adapted from Ref.⁷²

References

- (1) Chao, Q.; Ding, Y.; Chen, Z.-H.; Xiang, M.-H.; Wang, N.; Gao, X.-D. Recent Progress in Chemo-Enzymatic Methods for the Synthesis of N-Glycans. *Frontiers in Chemistry* **2020**, *8*.
- (2) Breton, C.; Šnajdrová, L.; Jeanneau, C.; Koča, J.; Imberty, A. Structures and Mechanisms of Glycosyltransferases. *Glycobiology* **2006**, *16* (2), 29R-37R.
<https://doi.org/10.1093/glycob/cwj016>.
- (3) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. Monosaccharide Metabolism. In *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press, 1999.
- (4) Freeze, H. H.; Elbein, A. D. Glycosylation Precursors. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009.
- (5) Varki, A.; Kornfeld, S. Historical Background and Overview. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.
- (6) Schneider, M.; Al-Shareffi, E.; Haltiwanger, R. S. Biological Functions of Fucose in Mammals. *Glycobiology* **2017**, *27* (7), 601–618. <https://doi.org/10.1093/glycob/cwx034>.
- (7) Varki, A. Sialic Acids in Human Health and Disease. *Trends Mol Med* **2008**, *14* (8), 351–360.
<https://doi.org/10.1016/j.molmed.2008.06.002>.
- (8) Xu, G.; Wong, M.; Li, Q.; Park, D.; Cheng, Z.; Lebrilla, C. B. Unveiling the Metabolic Fate of Monosaccharides in Cell Membranes with Glycomic and Glycoproteomic Analyses. *Chem Sci* **2019**, *10* (29), 6992–7002. <https://doi.org/10.1039/c9sc01653h>.
- (9) Altman, M. O.; Gagneux, P. Absence of Neu5Gc and Presence of Anti-Neu5Gc Antibodies in Humans—An Evolutionary Perspective. *Front Immunol* **2019**, *10*, 789.
<https://doi.org/10.3389/fimmu.2019.00789>.
- (10) Peri, S.; Kulkarni, A.; Feyertag, F.; Berninsone, P. M.; Alvarez-Ponce, D. Phylogenetic Distribution of CMP-Neu5Ac Hydroxylase (CMAH), the Enzyme Synthetizing the

- Proinflammatory Human Xenoantigen Neu5Gc. *Genome Biol Evol* **2017**, *10* (1), 207–219.
<https://doi.org/10.1093/gbe/evx251>.
- (11) Suzuki, T.; Cummings, R. D.; Aebi, M.; Parodi, A. Glycans in Glycoprotein Quality Control. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.
- (12) Schnaar, R. L.; Sandhoff, R.; Tiemeyer, M.; Kinoshita, T. Glycosphingolipids. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.
- (13) Stanley, P.; Moremen, K. W.; Lewis, N. E.; Taniguchi, N.; Aebi, M. N-Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.
- (14) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. O-Glycans. In *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press, 1999.
- (15) Zhang, L.; Li, Y.; Li, R.; Yang, X.; Zheng, Z.; Fu, J.; Yu, H.; Chen, X. Glycoprotein In Vitro N-Glycan Processing Using Enzymes Expressed in E. Coli. *Molecules* **2023**, *28* (6), 2753.
<https://doi.org/10.3390/molecules28062753>.
- (16) Li, Q.; Xie, Y.; Wong, M.; Barboza, M.; Lebrilla, C. B. Comprehensive Structural Glycomic Characterization of the Glycocalyxes of Cells and Tissues. *Nat Protoc* **2020**, *15* (8), 2668–2704. <https://doi.org/10.1038/s41596-020-0350-4>.
- (17) Aguilar, A. L.; Briard, J. G.; Yang, L.; Macauley, M. S.; Wu, P. Tools for Studying Glycans: Recent Advances in Chemoenzymatic Glycan Labeling. *ACS Chem Biol* **2017**, *12* (3), 611–621. <https://doi.org/10.1021/acscchembio.6b01089>.
- (18) Shah, N.; Kuntz, D. A.; Rose, D. R. Comparison of Kifunensine and 1-Deoxymannojirimycin Binding to Class I and II Alpha-Mannosidases Demonstrates Different Saccharide Distortions in Inverting and Retaining Catalytic Mechanisms. *Biochemistry* **2003**, *42* (47), 13812–13816. <https://doi.org/10.1021/bi034742r>.

- (19) Legler, K.; Rosprim, R.; Karius, T.; Eylmann, K.; Rossberg, M.; Wirtz, R. M.; Müller, V.; Witzel, I.; Schmalfeldt, B.; Milde-Langosch, K.; Oliveira-Ferrer, L. Reduced Mannosidase MAN1A1 Expression Leads to Aberrant N-Glycosylation and Impaired Survival in Breast Cancer. *Br J Cancer* **2018**, *118* (6), 847–856. <https://doi.org/10.1038/bjc.2017.472>.
- (20) Sheng, Y.; Vinjamuri, A.; Alvarez, M. R. S.; Xie, Y.; McGrath, M.; Chen, S.; Barboza, M.; Frieman, M.; Lebrilla, C. B. Host Cell Glycocalyx Remodeling Reveals SARS-CoV-2 Spike Protein Glycomic Binding Sites. *Front Mol Biosci* **2022**, *9*, 799703. <https://doi.org/10.3389/fmolb.2022.799703>.
- (21) Zhou, Y.; Fukuda, T.; Hang, Q.; Hou, S.; Isaji, T.; Kameyama, A.; Gu, J. Inhibition of Fucosylation by 2-Fluorofucose Suppresses Human Liver Cancer HepG2 Cell Proliferation and Migration as Well as Tumor Formation. *Sci Rep* **2017**, *7* (1), 11563. <https://doi.org/10.1038/s41598-017-11911-9>.
- (22) Zhou, J. Y.; Oswald, D. M.; Oliva, K. D.; Kreisman, L. S. C.; Cobb, B. A. The Glycoscience of Immunity. *Trends Immunol* **2018**, *39* (7), 523–535. <https://doi.org/10.1016/j.it.2018.04.004>.
- (23) Trzos, S.; Link-Lenczowski, P.; Pocheć, E. The Role of N-Glycosylation in B-Cell Biology and IgG Activity. The Aspects of Autoimmunity and Anti-Inflammatory Therapy. *Front Immunol* **2023**, *14*, 1188838. <https://doi.org/10.3389/fimmu.2023.1188838>.
- (24) Vicente, M. M.; Leite-Gomes, E.; Pinho, S. S. Glycome Dynamics in T and B Cell Development: Basic Immunological Mechanisms and Clinical Applications. *Trends Immunol* **2023**, *44* (8), 585–597. <https://doi.org/10.1016/j.it.2023.06.004>.
- (25) Pinho, S. S.; Alves, I.; Gaifem, J.; Rabinovich, G. A. Immune Regulatory Networks Coordinated by Glycans and Glycan-Binding Proteins in Autoimmunity and Infection. *Cell Mol Immunol* **2023**, *20* (10), 1101–1113. <https://doi.org/10.1038/s41423-023-01074-1>.
- (26) Krautler, F.; Iqbal, A. J. Glycans and Glycan-Binding Proteins as Regulators and Potential Targets in Leukocyte Recruitment. *Front Cell Dev Biol* **2021**, *9*, 624082. <https://doi.org/10.3389/fcell.2021.624082>.

- (27) de Haan, N.; Falck, D.; Wuhler, M. Monitoring of Immunoglobulin N- and O-Glycosylation in Health and Disease. *Glycobiology* **2019**, *30* (4), 226–240.
<https://doi.org/10.1093/glycob/cwz048>.
- (28) Higel, F.; Seidl, A.; Sörgel, F.; Friess, W. N-Glycosylation Heterogeneity and the Influence on Structure, Function and Pharmacokinetics of Monoclonal Antibodies and Fc Fusion Proteins. *European Journal of Pharmaceutics and Biopharmaceutics* **2016**, *100*, 94–100.
<https://doi.org/10.1016/j.ejpb.2016.01.005>.
- (29) Maverakis, E.; Kim, K.; Shimoda, M.; Gershwin, M. E.; Patel, F.; Wilken, R.; Raychaudhuri, S.; Ruhaak, L. R.; Lebrilla, C. B. Glycans in the Immune System and The Altered Glycan Theory of Autoimmunity: A Critical Review. *Journal of Autoimmunity* **2015**, *57*, 1–13.
<https://doi.org/10.1016/j.jaut.2014.12.002>.
- (30) Ząbczyńska, M.; Polak, K.; Kozłowska, K.; Sokołowski, G.; Pocheć, E. The Contribution of IgG Glycosylation to Antibody-Dependent Cell-Mediated Cytotoxicity (ADCC) and Complement-Dependent Cytotoxicity (CDC) in Hashimoto's Thyroiditis: An in Vitro Model of Thyroid Autoimmunity. *Biomolecules* **2020**, *10* (2), 171. <https://doi.org/10.3390/biom10020171>.
- (31) Ruhaak, L. R.; Barkauskas, D. A.; Torres, J.; Cooke, C. L.; Wu, L. D.; Stroble, C.; Ozcan, S.; Williams, C. C.; Camorlinga, M.; Rocke, D. M.; Lebrilla, C. B.; Solnick, J. V. The Serum Immunoglobulin G Glycosylation Signature of Gastric Cancer. *EuPA Open Proteom* **2015**, *6*, 1–9. <https://doi.org/10.1016/j.euprot.2014.11.002>.
- (32) Alvarez, M. R. S.; Zhou, Q.; Tena, J.; Lebrilla, C. B.; Completo, G. C.; Heralde, F. M. I.; Cabanatan, M.; Barzaga, Ma. T.; Tan-Liu, N.; Ladrera, G. I.; Danguilan, J. L.; Rabajante, J.; Padolina, I.; Nacario, R. C. N-Glycan and Glycopeptide Serum Biomarkers in Philippine Lung Cancer Patients Identified Using Liquid Chromatography–Tandem Mass Spectrometry. *ACS Omega* **2022**, *7* (44), 40230–40240. <https://doi.org/10.1021/acsomega.2c05111>.
- (33) Tena, J.; Tang, X.; Zhou, Q.; Harvey, D.; Barajas-Mendoza, M.; Jin, L.-W.; Maezawa, I.; Zivkovic, A. M.; Lebrilla, C. B. Glycosylation Alterations in Serum of Alzheimer's Disease Patients Show Widespread Changes in N-Glycosylation of Proteins Related to Immune Function, Inflammation, and Lipoprotein Metabolism. *Alzheimers Dement (Amst)* **2022**, *14* (1), e12309. <https://doi.org/10.1002/dad2.12309>.

- (34) Matsumoto, A.; Shikata, K.; Takeuchi, F.; Kojima, N.; Mizuochi, T. Autoantibody Activity of IgG Rheumatoid Factor Increases with Decreasing Levels of Galactosylation and Sialylation¹. *The Journal of Biochemistry* **2000**, *128* (4), 621–628. <https://doi.org/10.1093/oxfordjournals.jbchem.a022794>.
- (35) Glycan Profiling of Anti–Citruillinated Protein Antibodies Isolated from Human Serum and Synovial Fluid. <https://doi.org/10.1002/art.27414>.
- (36) Novak, J.; Moldoveanu, Z.; Renfrow, M. B.; Yanagihara, T.; Suzuki, H.; Raska, M.; Hall, S.; Brown, R.; Huang, W.-Q.; Goepfert, A.; Kilian, M.; Poulsen, K.; Tomana, M.; Wyatt, R. J.; Julian, B. A.; Mestecky, J. IgA Nephropathy and Henoch-Schoenlein Purpura Nephritis: Aberrant Glycosylation of IgA1, Formation of IgA1-Containing Immune Complexes, and Activation of Mesangial Cells. **2007**. <https://doi.org/10.1159/000102455>.
- (37) Mestecky, J.; Tomana, M.; Moldoveanu, Z.; Julian, B. A.; Suzuki, H.; Matousovic, K.; Renfrow, M. B.; Novak, L.; Wyatt, R. J.; Novak, J. Role of Aberrant Glycosylation of IgA1 Molecules in the Pathogenesis of IgA Nephropathy. *Kidney Blood Press Res* **2008**, *31* (1), 29–37. <https://doi.org/10.1159/000112922>.
- (38) Mestecky, J.; Tomana, M.; Crowley-Nowick, P. A.; Moldoveanu, Z.; Julian, B. A.; Jackson, S. Defective Galactosylation and Clearance of IgA1 Molecules as a Possible Etiopathogenic Factor in IgA Nephropathy¹. **1993**. <https://doi.org/10.1159/000422410>.
- (39) Allen, A. C.; Harper, S. J.; Feehally, J. Galactosylation of N- and O-Linked Carbohydrate Moieties of IgA1 and IgG in IgA Nephropathy. *Clin Exp Immunol* **1995**, *100* (3), 470–474. <https://doi.org/10.1111/j.1365-2249.1995.tb03724.x>.
- (40) Ackerman, M. E.; Crispin, M.; Yu, X.; Baruah, K.; Boesch, A. W.; Harvey, D. J.; Dugast, A.-S.; Heizen, E. L.; Ercan, A.; Choi, I.; Streeck, H.; Nigrovic, P. A.; Bailey-Kellogg, C.; Scanlan, C.; Alter, G. Natural Variation in Fc Glycosylation of HIV-Specific Antibodies Impacts Antiviral Activity. *J Clin Invest* **2013**, *123* (5), 2183–2192. <https://doi.org/10.1172/JCI65708>.
- (41) Selman, M. H. J.; Jong, S. E. de; Soonawala, D.; Kroon, F. P.; Adegnika, A. A.; Deelder, A. M.; Hokke, C. H.; Yazdanbakhsh, M.; Wuhler, M. Changes in Antigen-Specific IgG1 Fc N-Glycosylation Upon Influenza and Tetanus Vaccination. *Molecular & Cellular Proteomics* **2012**, *11* (4). <https://doi.org/10.1074/mcp.M111.014563>.

- (42) Selman, M. H. J.; Niks, E. H.; Titulaer, M. J.; Verschuuren, J. J. G. M.; Wuhrer, M.; Deelder, A. M. IgG Fc N-Glycosylation Changes in Lambert-Eaton Myasthenic Syndrome and Myasthenia Gravis. *J. Proteome Res.* **2011**, *10* (1), 143–152. <https://doi.org/10.1021/pr1004373>.
- (43) Holland, M.; Yagi, H.; Takahashi, N.; Kato, K.; Savage, C. O. S.; Goodall, D. M.; Jefferis, R. Differential Glycosylation of Polyclonal IgG, IgG-Fc and IgG-Fab Isolated from the Sera of Patients with ANCA-Associated Systemic Vasculitis. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2006**, *1760* (4), 669–677. <https://doi.org/10.1016/j.bbagen.2005.11.021>.
- (44) HOLLAND, M.; TAKADA, K.; OKUMOTO, T.; TAKAHASHI, N.; KATO, K.; ADU, D.; BEN-SMITH, A.; HARPER, L.; SAVAGE, C. O. S.; JEFFERIS, R. Hypogalactosylation of Serum IgG in Patients with ANCA-Associated Systemic Vasculitis. *Clinical and Experimental Immunology* **2002**, *129* (1), 183–190. <https://doi.org/10.1046/j.1365-2249.2002.01864.x>.
- (45) Saldova, R.; Royle, L.; Radcliffe, C. M.; Abd Hamid, U. M.; Evans, R.; Arnold, J. N.; Banks, R. E.; Hutson, R.; Harvey, D. J.; Antrobus, R.; Petrescu, S. M.; Dwek, R. A.; Rudd, P. M. Ovarian Cancer Is Associated with Changes in Glycosylation in Both Acute-Phase Proteins and IgG. *Glycobiology* **2007**, *17* (12), 1344–1356. <https://doi.org/10.1093/glycob/cwm100>.
- (46) Espy, C.; Morelle, W.; Kavian, N.; Grange, P.; Goulvestre, C.; Viallon, V.; Chéreau, C.; Pagnoux, C.; Michalski, J.-C.; Guillevin, L.; Weill, B.; Batteux, F.; Guilpain, P. Sialylation Levels of Anti-Proteinase 3 Antibodies Are Associated with the Activity of Granulomatosis with Polyangiitis (Wegener's). *Arthritis Rheum* **2011**, *63* (7), 2105–2115. <https://doi.org/10.1002/art.30362>.
- (47) Youinou, P.; Pennec, Y.-L.; Casburn-Budd, R.; Dueymes, M.; Letoux, G.; Lamour, A. Galactose Terminating Oligosaccharides of IgG in Patients with Primary Sjögren's Syndrome. *Journal of Autoimmunity* **1992**, *5* (3), 393–400. [https://doi.org/10.1016/0896-8411\(92\)90151-F](https://doi.org/10.1016/0896-8411(92)90151-F).
- (48) Tretter, V.; Altmann, F.; März, L. 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* **1991**, *199* (3), 647–652. <https://doi.org/10.1111/j.1432-1033.1991.tb16166.x>.
- (49) Kawasaki, N. N-Glycanase Digestion (GlycoPOD), 2016. <https://jcggdb.jp/GlycoPOD>.

- (50) Müller, T.; Winter, D. Systematic Evaluation of Protein Reduction and Alkylation Reveals Massive Unspecific Side Effects by Iodine-Containing Reagents. *Molecular & Cellular Proteomics : MCP* **2017**, *16* (7), 1173. <https://doi.org/10.1074/mcp.M116.064048>.
- (51) Manea, M.; Mezo, G.; Hudecz, F.; Przybylski, M. Mass Spectrometric Identification of the Trypsin Cleavage Pathway in Lysyl-Proline Containing Oligopeptides. *J Pept Sci* **2007**, *13* (4), 227–236. <https://doi.org/10.1002/psc.836>.
- (52) Jakoby, T.; van den Berg, B. H.; Tholey, A. Quantitative Protease Cleavage Site Profiling Using Tandem-Mass-Tag Labeling and LC-MALDI-TOF/TOF MS/MS Analysis. *J Proteome Res* **2012**, *11* (3), 1812–1820. <https://doi.org/10.1021/pr201051e>.
- (53) Coskun, O. Separation Techniques: Chromatography. *Northern Clinics of Istanbul* **2016**, *3* (2), 156. <https://doi.org/10.14744/nci.2016.32757>.
- (54) *Gas Chromatography - an overview | ScienceDirect Topics*.
<https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/gas-chromatography> (accessed 2024-03-12).
- (55) Pitt, J. J. Principles and Applications of Liquid Chromatography-Mass Spectrometry in Clinical Biochemistry. *The Clinical Biochemist Reviews* **2009**, *30* (1), 19.
- (56) Jandera, P. LIQUID CHROMATOGRAPHY | Normal Phase☆. In *Encyclopedia of Analytical Science (Third Edition)*; Worsfold, P., Poole, C., Townshend, A., Miró, M., Eds.; Academic Press: Oxford, 2013; pp 162–173. <https://doi.org/10.1016/B978-0-12-409547-2.00300-0>.
- (57) *Reversed Phase Liquid Chromatography - an overview | ScienceDirect Topics*.
<https://www.sciencedirect.com/topics/nursing-and-health-professions/reversed-phase-liquid-chromatography> (accessed 2024-03-12).
- (58) Zhu, R.; Huang, Y.; Zhao, J.; Zhong, J.; Mechref, Y. Isomeric Separation of N-Glycopeptides Derived from Glycoproteins by Porous Graphitic Carbon (PGC) LC-MS/MS. *Anal Chem* **2020**, *92* (14), 9556–9565. <https://doi.org/10.1021/acs.analchem.0c00668>.
- (59) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry* **2012**, *2012*. <https://doi.org/10.1155/2012/282574>.

- (60) *Time-of-Flight Mass Spectrometry - an overview | ScienceDirect Topics*.
<https://www.sciencedirect.com/topics/materials-science/time-of-flight-mass-spectrometry>
(accessed 2024-03-12).
- (61) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **2013**, *85* (11), 5288–5296. <https://doi.org/10.1021/ac4001223>.
- (62) *Ion Trapping - an overview | ScienceDirect Topics*.
<https://www.sciencedirect.com/topics/nursing-and-health-professions/ion-trapping>
(accessed 2024-03-12).
- (63) Guo, Z.; Pang, J.; Liu, Z. On the Journey Exploring Nanoscale Packing Materials for Ultra-Efficient Liquid Chromatographic Separation. *Journal of Chromatography Open* **2022**, *2*, 100033. <https://doi.org/10.1016/j.jcoa.2022.100033>.
- (64) Ashwood, C.; Pratt, B.; MacLean, B. X.; Gundry, R. L.; Packer, N. H. Standardization of PGC-LC-MS-Based Glycomics for Sample Specific Glycotyping. *Analyst* **2019**, *144* (11), 3601–3612. <https://doi.org/10.1039/c9an00486f>.
- (65) West, C.; Elfakir, C.; Lafosse, M. Porous Graphitic Carbon: A Versatile Stationary Phase for Liquid Chromatography. *J Chromatogr A* **2010**, *1217* (19), 3201–3216.
<https://doi.org/10.1016/j.chroma.2009.09.052>.
- (66) Heil, L. R.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Hsu, C.; Searle, B. C.; Shulman, N.; Riffle, M.; Connolly, B.; MacLean, B. X.; Remes, P. M.; Senko, M. W.; Stewart, H. I.; Hock, C.; Makarov, A. A.; Hermanson, D.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data Independent Acquisition. *bioRxiv*.
<https://doi.org/10.1101/2023.06.03.543570>.
- (67) *Bottom-up Proteomics - an overview | ScienceDirect Topics*.
<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bottom-up-proteomics> (accessed 2024-03-12).
- (68) Olsen, J. V.; Ong, S.-E.; Mann, M. Trypsin Cleaves Exclusively C-Terminal to Arginine and Lysine Residues*. *Molecular & Cellular Proteomics* **2004**, *3* (6), 608–614.
<https://doi.org/10.1074/mcp.T400003-MCP200>.

- (69) Macias, L. A.; Santos, I. C.; Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal Chem* **2020**, 92 (1), 227–251. <https://doi.org/10.1021/acs.analchem.9b04859>.
- (70) Hao, Z.; Hong, Q.; Zhang, F.; Wu, S.-L.; Bennett, P. Current Methods for the Characterization of Posttranslational Modifications in Therapeutic Proteins Using Orbitrap Mass Spectrometry. In *Protein Analysis using Mass Spectrometry*; John Wiley & Sons, Ltd, 2017; pp 21–34. <https://doi.org/10.1002/9781119371779.ch3>.
- (71) Brun, V.; Dupuis, A.; Adrait, A.; Marcellin, M.; Thomas, D.; Court, M.; Vandenesch, F.; Garin, J. Isotope-Labeled Protein Standards: Toward Absolute Quantitative Proteomics. *Mol Cell Proteomics* **2007**, 6 (12), 2139–2149. <https://doi.org/10.1074/mcp.M700163-MCP200>.
- (72) Kim, B. J.; Dallas, D. C. Systematic Examination of Protein Extraction, Proteolytic Glycopeptide Enrichment and MS/MS Fragmentation Techniques for Site-Specific Profiling of Human Milk N-Glycoproteins. *Talanta* **2021**, 224, 121811. <https://doi.org/10.1016/j.talanta.2020.121811>.

Chapter II

Investigating The Effects of Glycosylation Inhibitors and Monosaccharides on Cell Surface and Secreted Antibody (IgG1) Glycosylation of CESS Lymphoblast Cell Line Using LC- MS/MS

Authors: Armin Oloumi¹, Michael Russelle Alvarez¹, Ryan L. Schindler¹, Qingwen Zhou¹, Ying Sheng¹, Yasmine Bouchibti¹, Yiyun Liu¹, Sheryl Grijaldo Alvarez¹, Emanuel Maverakis², Carlito Lebrilla¹

1. Department of Chemistry, University of California, Davis, CA
2. Department of Dermatology, University of California, Davis School of Medicine, Sacramento, CA

Abstract

Cell surface and secreted antibody glycans play vital roles in biological functions of immune cells, but the link between their cell surface glycosylation and secreted glycoprotein glycosylation from the same cells remains unexplored. In this study we demonstrated the effects of different glycosylation inhibitors on the glycosylation of the cell surface of immune cells (CESS) and the secreted antibodies (IgG1) from the same cells using nano LC-MS/MS based glycomics and glycoproteomics. Cell surface glycans and secreted IgG1 glycans were changed drastically after a 5-day treatment with enzyme inhibitors, such as those inhibiting fucosylation, sialylation and other structural features. Changes observed in glycosylation of cell surface mirrored the changes observed in glycosylation of secreted antibodies, as they were both selectively modified with the inhibitors. Additionally, proliferation assays and transcriptomics analysis were performed on the cells to assess viability and functionality of the cells after the treatments. The induced changes in glycans had no significant impact on viability and overall transcriptomics of the cells. Additions of galactose, fucose and fructose monosaccharides, which can act as precursors or building blocks for glycans, were also tested for induced changes in glycan structures. Supplementation of the cell medium with monosaccharides caused minor but significant changes in both cell surface and secreted antibody glycosylation. A sialylated glycan structure was only observed on secreted IgG1 at Asn 297 with supplementation of any of three monosaccharides. In conclusion, metabolic engineering tools for glycosylation are effective at changing both cell surface and secreted antibody glycans and provide promising potentials for immune cell targeting or antibody-based therapeutics.

1. Introduction

The glycocalyx is a protective and functional layer of carbohydrates that covers the surface of many types of cells, including blood cells.¹ One of the major components of the glycocalyx are N-glycans. N-Glycans are essential PTMs of glycoproteins that exert profound influences on protein function and, in turn, the respective biological processes.² N-Glycans are attached to asparagine residues on glycoproteins and consist of a common core of N-acetylglucosamine (GlcNAc) and mannose residues.³ Functional and structural diversity of N-glycans arise from their non-template driven biosynthesis involving various enzymes that assemble the monosaccharide building blocks into precise structures with unique compositions and branching.⁴

N-Glycans are often classified into the three major groups corresponding to high-mannose, complex, and hybrid-type structures.⁵ High mannose structures are synthesized at an earlier stage in the process in the ER, whereas complex and hybrid-type N-glycans are formed subsequently by further processing with glycoenzymes in the Golgi.⁶ These enzymes are involved in trimming and extending the nascent glycans in process such as galactosylation, fucosylation, and sialylation. Galactosylation, the addition of galactose, contributes to glycoprotein stability⁷ and modulation of immune responses.⁸ On the other hand, fucosylation, the addition of fucose, plays a crucial role in cell adhesion,⁹ inflammation,¹⁰ and immune recognition.¹¹ Sialylation, the addition of sialic acid residues, regulates glycoprotein interactions with other proteins, impacting cellular signaling,¹² immune evasion,¹³ and clearance.¹⁴ These diverse glycan features underscore the importance of N-glycan structures in orchestrating a wide range of biological processes and cellular functions. Therefore, inhibition of any of these steps would contribute to changes in structure and hence biological functions of the respective proteins. A clear example of this was

demonstrated in our previous study where cell-cell adhesion of Caco-2 cells was reduced with inhibition of fucosylation.⁹

Glycans on the immune cell surface similarly affect the function of immune cells and greatly influence secreted immunoglobulins, which are antibodies produced by the immune cells.¹⁵ For example, the presence (or absence) of fucose residues on the N-glycans of the Fc region of antibodies greatly affect their ability to bind to the Fc receptors on immune cells.¹⁶ Altered glycosylation can also enhance or inhibit antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP).¹⁷ Previous studies from our laboratory and elsewhere have demonstrated that changes in IgG glycosylation in sera is associated with gastric cancer,¹⁸ lung cancer,¹⁹ and Alzheimer's disease patients.²⁰ Similarly, IgG glycosylation changes have also been observed in COVID-19 patients.^{21,22}

Despite the significant roles of N-glycans on cell surface and secreted antibodies, the link between cell surface glycosylation of immune cells and their secreted antibodies remains unexplored. The ability to selectively modify these glycans has significant therapeutic applications and may be used in production of the most efficient proteoforms of therapeutic antibodies. As the modification of the glycans involve the addition of monosaccharide to the cells, the use of selective monosaccharide supplements would also provide a greater understanding of the effects of supplements and diet treatments on the immune system. Indeed, we have previously explored this concept by incorporating intact C13 labeled monosaccharides, such as fucose, into the membrane glycans of epithelial cell *in vitro*.²⁴

The tools for studying the impact of specific glycan structures are often limited to genetic²⁵ and metabolic engineering,²⁶ however it has generally been difficult to characterize the induced

changes in specific structures. Advances in nanoflow liquid chromatography (nanoLC) and tandem mass spectrometry have paved the way for further advances in glycomics and glycoproteomics.²⁷ We have demonstrated the utility of these advances by developing methods that enable the analysis of released N-glycans from the cell surfaces while providing both identifications and abundances of individual glycan structures.²⁸ Additionally, we have developed methods that allow the identification and quantitation of glycopeptides on antibodies in a site-specific manner.²⁹ The combined use of these methods allows us to link the glycosylation of cell surface membrane proteins to secreted antibodies. Membrane glycoproteins and secreted antibodies are glycosylated in the same cellular organelles, but the glycoprotein transfers are performed after glycan biosynthesis, and changes in glycosylation can potentially alter the glycoprotein transfer in their respective pathways.

In this study, we aim to unravel the link between cell surface glycosylation and secreted antibody glycosylation and demonstrate how metabolic altering techniques such as glycosylation inhibition or monosaccharide supplementations can be used to modulate or alter glycosylation on the cell membrane and on secreted proteins. We further examine how glycans produced on secreted IgG1 are affected and whether the glycoprotein functions are altered. Additionally, we monitored the effects of glycosylation inhibitors on the cell transcriptome through 3'-tag poly tail RNA analysis.

2. Materials and Methods

2.1. Cell Culture and Treatments

CESS cells were cultured according to the protocols provided by the American Type Culture Collection (ATCC) using RPMI-1640 (ATCC, cat. 30-2001) growth medium supplemented with 10% (v/v) FBS (GIBCO, cat. 26140095) and 1% (v/v) penicillin–streptomycin (GIBCO, cat. 15140-1220) in a cell culture incubator with a humidified atmosphere containing 5% CO₂ at 37 °C. The cells were maintained at a population of less than 1×10^6 cells/mL and were used after 5 passages. Glycosylation inhibitor treatments with kifunensine (mannosidase inhibitor), 2-deoxy-2-fluoro-l-fucose (2-F-Fucose, fucosylation inhibitor), and 2,4,7,8,9-Penta-O-acetyl-N-acetyl-3-fluoro-b-d-neuraminic acid methyl ester (3-F-Sia, sialylation inhibitor) (Carbosynth, San Diego, CA, USA) were initiated at 2×10^5 cells/mL confluency. Thereafter, the medium was changed to AIM-V (GIBCO, cat. 12055083) with continuation of the treatments for an additional two days when the cells had reached 9×10^5 cells/mL confluency. The cells were harvested afterwards, and the media was stored in -80 °C prior to IgG purification.

2.2 Cell Surface Glycomics

The Glycomics analysis was performed based on a previously published protocol.²⁸ Briefly, the cells were centrifuged at 1,000 g for 3 min at 4 °C and they were washed with 20 mL of D-PBS twice prior to reconstitution in 1.2 mL of 0.25 M sucrose (MilliporeSigma, cat. S7903), 20mM HEPES (Thermo Fisher Scientific, cat. 15630080) adjusted to pH 7.4 with KOH (MilliporeSigma, cat. P5958) and a 1:100 protease inhibitor cocktail (Calbio-chem, cat. 539137).

The cells were lysed by sonication for 1 minute using probe-tip (QSonica, cat. Q700-110), which was set to the amplitude of 25 with 5 s 'on' and 10 s 'off' cycle. The cells were then centrifuged at 2,000 g for 10 minutes at 4 °C to remove the cell debris, and the supernatant was ultracentrifuged at 200,000 g for 45 minutes at 4°C to pull down the cell membrane. The membrane was resuspended in 0.5 mL of 0.2 M Na₂CO₃ (MilliporeSigma, cat. P5958) and ultracentrifuged again to remove membrane-associated proteins. The last ultracentrifugation was performed in 0.5 mL of water to remove the salts.

2.3. Sample Preparation for N-Glycan Analysis

The enriched membrane fraction was resuspended in 200 µL of 100 mM of ammonium bicarbonate (Sigma-Aldrich, cat. A6141) and 5 mM Dithiothreitol (Promega, cat. V3151) to denature the protein. The samples were placed in a 100°C water bath for two minutes with 10 second heating and cooling intervals to further denature the membrane proteins. After cooling the samples to room temperature, 2 µL of PNGase F (New England Biolabs, cat. P0705L) was added and the sample mixtures were incubated at 37°C for 18 hours to release the N-glycans. Samples were quenched with 350 µL of water and ultracentrifuged to remove the proteins from released N-glycans in solution. The supernatant containing the released N-glycans were cleaned using a porous graphitic carbon solid phase extraction plate (Glygen, cat. FNScar800). The wells were first washed with 80% (v/v) acetonitrile and 0.1% (v/v) trifluoroacetic acid in water and then conditioned with pure water. Samples were loaded onto the plate wells and washed with 600 µL of water three times. N-Glycans were then eluted with 600 µL of 80% (v/v)

acetonitrile and 0.1% (v/v) trifluoroacetic acid in water. Eluted N-glycans were dried and reconstituted in 30 μ L of pure water prior to analysis.

2.4 N-Glycan Analysis

N-Glycan analysis was performed on an Agilent 1200 series ChipCube system in positive ion mode and a 6520 Q-TOF mass spectrometer. N-Glycans were separated on a PGC-Chip II with 40 nL enrichment column (5 μ m, 75 mm $\times$ 43 mm, Agilent Technologies cat. G4240-62003). A binary gradient was applied using (A) 0.1% (v/v) formic acid and 3% (v/v) acetonitrile in water and (B) 1% formic acid and 90% (v/v) acetonitrile in water: 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). 5 μ L of each sample was injected, loaded onto the enrichment column, and washed for 10 minutes with 100% mobile phase B at 3 μ L/min. Next, the chip valve switched to analysis mode to move the enrichment column in line and the analytical gradient began with a flow rate of 0.33 μ L/min. Source conditions included the use of nitrogen as the drying gas with a temperature of 325°C flowing at 5 L/min and a capillary voltage of 1800 V. Auto MS/MS acquisition mode allowed four precursors per cycle with a threshold of 2000 counts and quadrupole filtering with a narrow isolation width for the 600–2000 m/z range. Collision-induced disassociation with nitrogen gas used a precursor-dependent formula $y=1.8*(m/z)/100-3.6$ to determine individual collision energies with a range of 100–2000 m/z. Continual mass correction with a reference calibration standard of 1221.990637 m/z was used (Agilent Technologies, cat. HP-1221). N-Glycan compound verification used

MassHunter Qual B08.00 software with the 'Find by Molecular Feature' algorithm and an in-house database containing N-glycan monoisotopic masses and descriptive information.

2.5 IgG Purification

The media was filtered using 100 K Da spin column (Amicon, cat. UFC910008), then the filtrant was subjected to IgG extraction using protein A/G magnetic beads (Pierce, cat. 88802) per manufacturer protocol. The extraction utilized 0.05% (v/v) Tween-20 for 20-fold dilution of the media and washing the beads, 0.1 M glycine (pH 2.0) solution for eluting the extracted IgG from the beads and 1 M Tris (pH 7.5) solution to neutralize the eluted IgG containing solution. The resulting extracted IgG solution was then buffer exchanged into 50 μ L of 50 mM ammonium bicarbonate (Sigma-Aldrich, cat. A6141) using 100 K Da spin column (Amicon, cat. UFC510008) to remove salts and prepare the protein for digestion in ammonium bicarbonate buffer.

2.6 IgG Digestion

Details of protein digestion for the glycoproteomics analysis have been described previously²⁸. The proteins were reduced with 2 μ L of 550 mM dithiothreitol (Promega, cat. V3151) for 1 hour at 60 °C and alkylated with 4 μ L of 450 mM iodoacetamide (Sigma-Aldrich, cat. I6125) in the dark for 30 minutes. The samples were incubated with 2 μ g of sequencing-grade trypsin (Promega, cat. V5111) at 37°C for 18 hours. The resulting peptides and glycopeptides were dried using a miVac (SP Scientific, PA, United States) prior to mass spectrometry analysis.

2.7 Glycoproteomic Analysis With LC-MS/MS

The digested 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, United States). The analytes were separated on an Acclaim™ PepMap™ 100 C18 LC Column (3 μ m, 0.075 mm $\times$ 150 mm, ThermoFisher Scientific, cat. 164946). A binary gradient was applied using 0.1% (v/v) formic acid in (A) water and (B) 80% (v/v) acetonitrile in water: 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 1.8 kV spray voltage, 275°C ion transfer capillary temperature, and the acquisition was performed with the full MS scanned from 700 to 2000 m/z 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.

2.8 Glycoproteomic Data Analysis

Glycopeptide fragmentation spectra were annotated using Byonic software (Protein Metrics, CA, United States) against the UniProt human proteome sequences. Common modifications, including cysteine carbamidomethyl, methionine oxidation, asparagine deamidation and glutamine deamidation were assigned. A published in-house N-glycan library was utilized for the glycopeptide identification, and relative abundance values were calculated based on the glycopeptide precursor peak areas.²⁸ Identified glycopeptides were filtered by 10

ppm mass filter accuracy and MS/MS spectra for every glycopeptide was checked manually to ensure accurate identification.

2.9 IgG1 Molecular Modeling and Molecular Dynamics

IgG1 in complex with the FCIIIR receptor (PDB ID: 3SGJ) was modeled to include N-glycans at FCIIIR receptor (3200) and IgG1-Asn 297 reflecting the structures observed in the inhibitor treatments: 4510 (control and sialic acid inhibitor), 9200 (kifunensine), 4500 (fucose inhibitor), and 5411 (monosaccharide supplemented). N-Glycan structures were added to the structure, and the resulting structure minimized using CHARMM-GUI Glycan modeler. The system was solvated using the TIP3P model, and 150 mM KCl was added. The CHARMM36mm force field was used for both the protein and glycans. Simulations were visualized, and the number of interacting hydrogen bonds between the glycans and proteins was analyzed using DiscoveryStudio (Dessault Systemes).

2.10 RNA extraction and 3'-TagSeq RNAseq analysis

Total RNA was extracted from cultured cells following the RNeasy Mini kit (Qiagen, cat. 74104) manufacturer instructions. Cells were harvested by trypsinization, then resuspended in buffer RLT to obtain cell lysates. These samples were further homogenized using the QiAshredder spin columns (Qiagen, cat. 79656), which were centrifuged at 21,000 xg for 2 minutes. To the flow-through, 70% ethanol solution was added, which were then transferred to the RNeasy spin column, centrifuging afterward at 8000 xg for 30 s. The spin columns were further washed with buffer RW1. To the spin columns, DNase I incubation mix (1:7 DNase I

solution in buffer RDD) was added, followed by incubation at room temperature for 15 minutes. The spin-columns were further washed with buffer RW1 by centrifuging at 8000 xg for 30 s. This was followed by twice centrifugation with buffer RPE at 8000 xg for 30 s and then 2 minutes. Finally, cleaned-up RNA samples were collected by adding sterile RNase-free H₂O to the spin-columns and centrifuging for 8000 xg for 1 minute. The samples were stored in -80 °C until submission to the UC Davis DNAtch Core Facility. The samples were analyzed using Batch 3' Tag-Seq analysis, with a read number of 4 M. Upon receipt of RNAseq data, raw reads were processed using a custom script based on the Tag-seq script (<https://github.com/ben-laufer/Tag-seq>) using the High-Performance Computing (HPC) cluster of the Philippine Department of Science and Technology. Data was processed using quantification mode to obtain normalized and annotated gene counts.

3. Results

Workflows for the cell surface glycomics and secreted IgG1 glycoproteomics of the CESS cell line are shown schematically in Figure 2.1. Cell surface glycomics analysis involved harvesting the cells via centrifugation, lysing and ultracentrifugation to enrich the cell membrane. The enriched membrane fractions were then subjected to N-glycan release using the enzyme PNGase F. The released oligosaccharides were enriched and analyzed using nanoLC–MS/MS with a PGC stationary phase and Q-TOF mass analyzer producing N-glycan chromatograms with isomer separation. Over 300 N-glycan structures are generally obtained

and identified with some cells having many more structures. The cell media after the growth were used for analysis of secreted antibodies. This glycoproteomic analysis involved concentrating the medium and enriching for IgG1 using protein A/G on magnetic beads, followed by tryptic digestion and nanoLC-MS/MS analysis using a C18 stationary phase coupled to an Orbitrap mass analyzer.

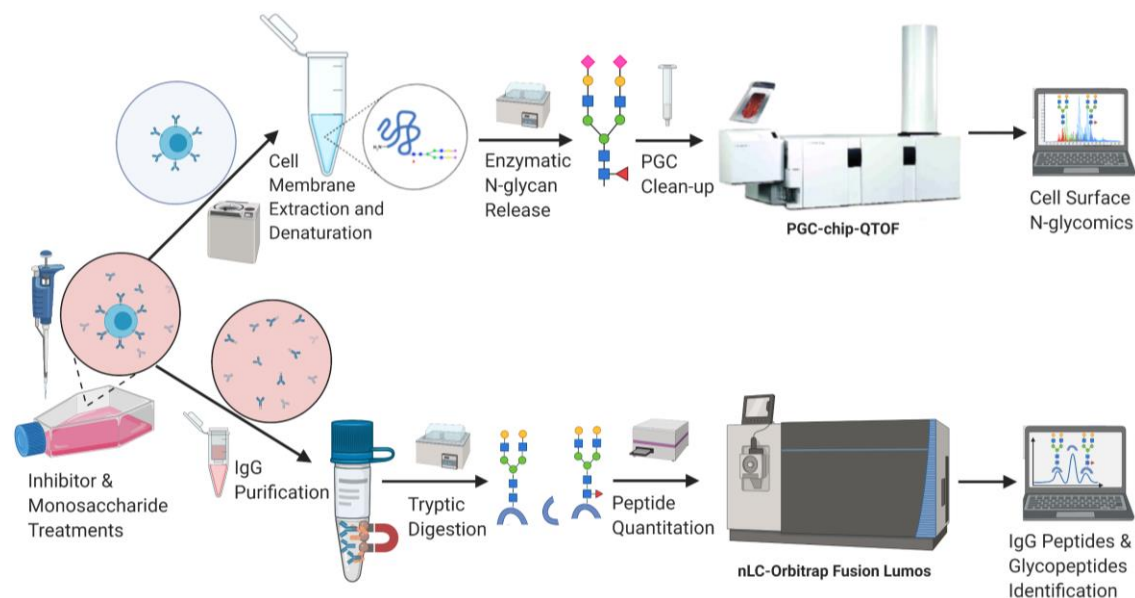


Figure 2.1. Schematics of Cell Surface Glycomics and Secreted Antibody Glycoproteomics workflow using nanoLC-MS/MS.

3.1. N-Glycomics of the Cell Membrane

3.1.1. The Effects of Glycosyl Transferase Inhibitors on the N-Glycome of the Cell Membrane

Cell surface glycomics analysis of CESS cell lines yielded identification of 607 glycan structures with the majority being sialofucosylated (60.1% total relative abundances, Supplementary Figure 2.1) containing both fucose and sialic acid residues. In addition to the

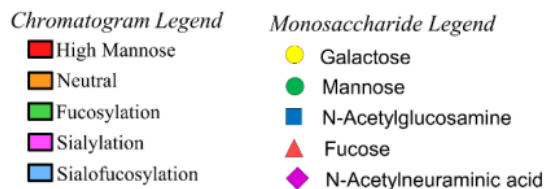
sialofucosylated N-glycans, other complex- or hybrid-type structures including fucosylated (17.0% relative abundance), sialylated (8.6% relative abundance), and undecorated or containing neither (3.3% relative abundance) and high-mannose (11.0% relative abundance) structures were observed (Supplementary Figure 2.1). Di-, tri- and tetra- antennary glycans as well as those containing bisecting GlcNAc (B-1,4 linked to core mannose) were among the identified structures (Figure 2.2a).

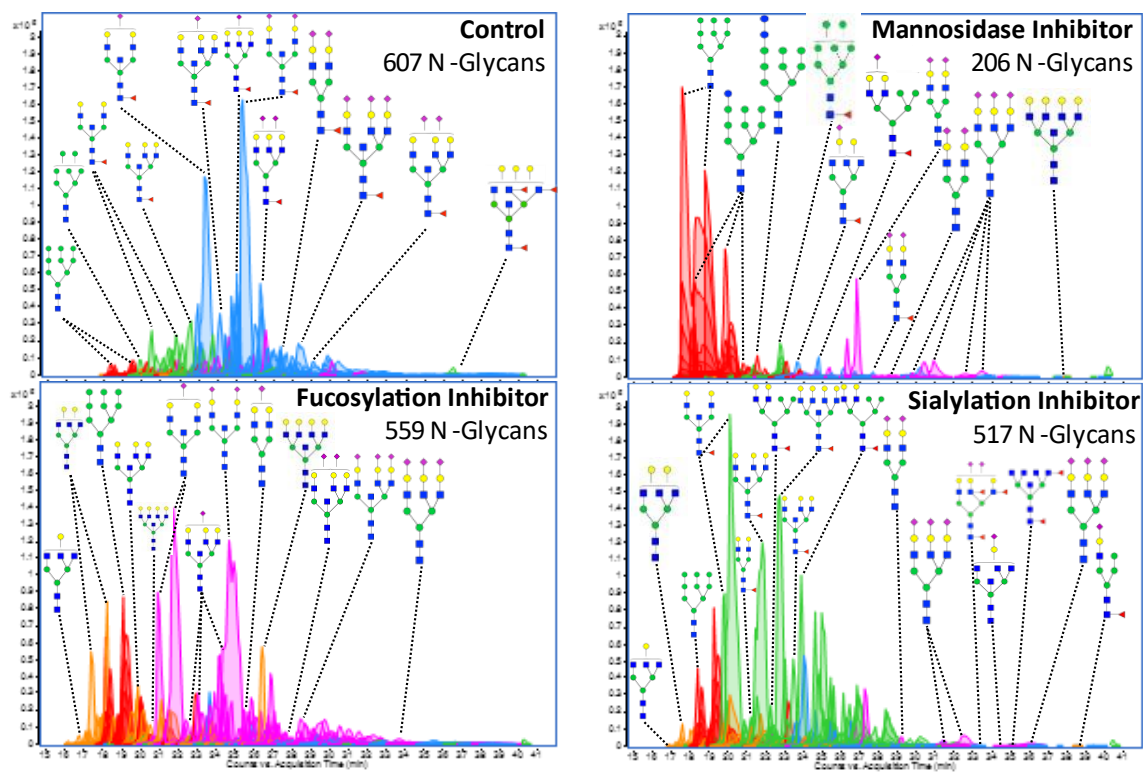
The effects of sialic acid inhibition were probed with the addition of 3-F-Sia (a fluorinated analogue of sialic acid) to the cell media. As the cell membrane of CESS contained mostly sialofucosylated glycans, treatment with the sialic acid inhibitor reduced the abundances of sialofucosylated species by 72% ($p=0.0007$, Figure 2.2b) while the nonsialylated but fucosylated glycan abundances increased by 570% ($P<0.0001$, Figure 2.2b). The results show that the structures that were previously sialofucosylated, were converted mainly to the corresponding asialylated fucosylated structures (Figure 2.2b).

The effects of inhibiting fucosylation were examined by adding 2-F-Fucose (a fluorinated analogue of fucose) to the cell media. The fucosylation inhibitor treatment reduced the abundances of sialofucosylated species by 82% ($p<0.0001$, Figure 2.2b), while the sialylated and undecorated glycan abundances increased by 1143% and 1639% respectively ($p<0.0001$ and $p=0.0012$, Figure 2.2b). The results indicate that the structures that were previously sialofucosylated, were converted to the corresponding afucosylated sialylated and undecorated structures (Figure 2.2b).

Kifunensine is an alkaloid with a high inhibitory effect on class I-mannosidases, such as MAN1A1, MAN1A2, MAN1B1, and MAN1C1.²⁸ These mannosidases trim mannose residues of Man-9, turning it into Man-8, Man-7, Man-6, and subsequently Man-5.²⁹ Man-8, Man-7, Man-6, and Man-5 have 8, 7, 6, and 5 mannose residues respectively. High mannose glycans are generally not fucosylated or sialylated, as the respective enzymes act later in the process. Therefore, inhibiting mannosidases with kifunensine often leads to production of mostly high-mannose glycans with no decorations.³⁰ Kifunensine treatment reduced the sialofucosylation glycan abundances by 91% ($p < 0.0001$, Figure 2.2b), and increased the abundances of high-mannose species by 1997% ($p < 0.0001$, Figure 2.2b). The results showed the glycans remained mostly high mannose type structures with 5 to 9 mannose residues (Figure 2.2a) without any GlcNAc, fucose and sialic acid residues.

a)





b)

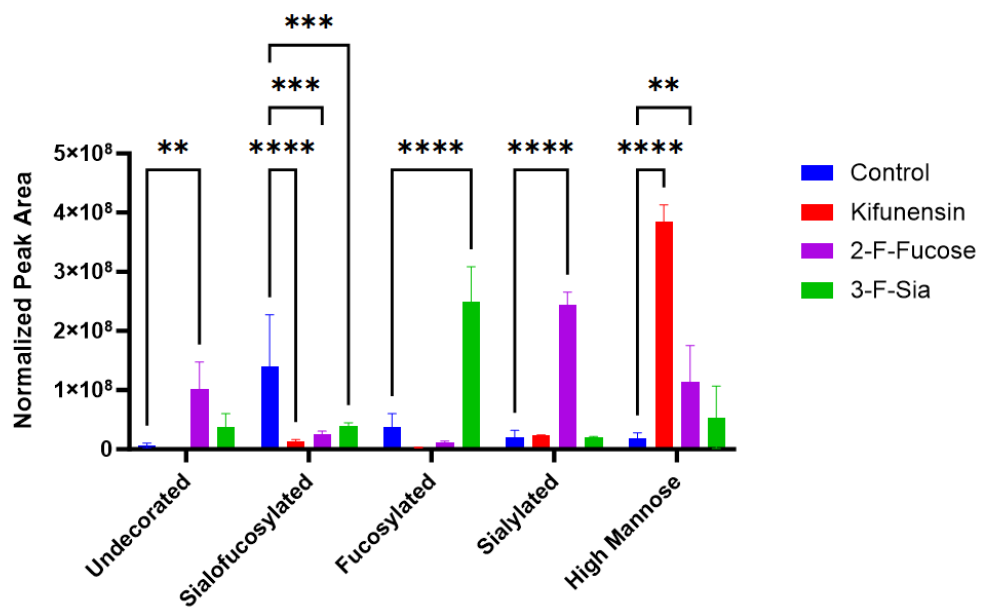


Figure 2.2. N-Glycome of cell membrane with glycosyl transferase inhibitors a) Chromatograms representing the separation and identification of the released cell membrane N-glycans in absence (control), and presence of the three inhibitors on a reversed phase nLC-MS/MS instrument. B) Bar graph of total abundances of types of identified structures based on their peak areas normalized to the membrane protein amounts.

3.1.2. The Effects of Monosaccharide Supplementations on the N-glycome of the Cell Membrane

Monosaccharides are a major constituent of human diet. It is therefore important to know how monosaccharides affect the glycosylation of immune cells and secreted glycoproteins. We sought to decipher these effects by supplementing the cells with additional monosaccharides, such as galactose, fructose and fucose. Upon transportation into cells via membrane transporters, monosaccharides undergo activation and direct incorporation into glycans. Alternatively, they may be converted into other monosaccharides via interconversion pathways.

Supplementation of cell media with galactose increased the abundances of sialofucosylated N-glycans by 96% ($p=0.0002$, Figure 2.3). While fructose supplementation increased both sialofucosylated and fucosylated type N-glycans abundances by 107% and 170% respectively ($p<0.0001$ and $p=0.03$ respectively, Figure 2.3). Furthermore, relative abundance comparisons of all identified N-glycans showed galactose supplementation increased the relative abundances of three N-glycans compared to the control, and two of these N-glycans were sialofucosylated type N-glycans (Supplementary Table 2.1).

There were no significant changes in specific glycan groups with the supplementation of cells with fucose (Figure 2.3). This may be due to multiple factors, such as the inability of the cells to import fucose from media into cytoplasm effectively. Additionally, there were no significant changes in relative abundances of the top 25 most abundant glycans observed in the samples, when compared to the control (Supplementary Figure 2.5). However, relative abundance comparisons of all identified N-glycans showed fucose supplementation increased the relative abundances of ten N-glycans compared to the control, and these ten N-glycans all contained fucose residues (Supplementary Table 2.1).

There was a significant increase in abundances of sialofucosylated glycans with supplementation of cells with fructose (Figure 2.3). This shows the cells utilized fructose to build and incorporate monosaccharides into the N-glycan structures. Additionally, an increase in fucosylated glycans was observed as well ($p=0.03$, Figure 2.3). Furthermore, relative abundance comparisons of all identified N-glycans showed fructose supplementation increased the relative abundances of four N-glycans and reduced the relative abundances of two N-glycans compared to the control (Supplementary Table 2.1). Two of the N-glycans that had increased relative abundances were fucosylated type N-glycans, while the other two with increased relative abundances were sialofucosylated type N-glycans (Supplementary Table 2.1).

Although the changes in abundances of glycan-types were significant in the supplementation of cells with galactose and fructose, they did not appear to heavily affect the production of top 25 most abundant glycans, as there was no significant change in relative abundances of the top

25 most abundant glycans observed in the control with respect to all identified glycans (Supplementary Figure 2.5).

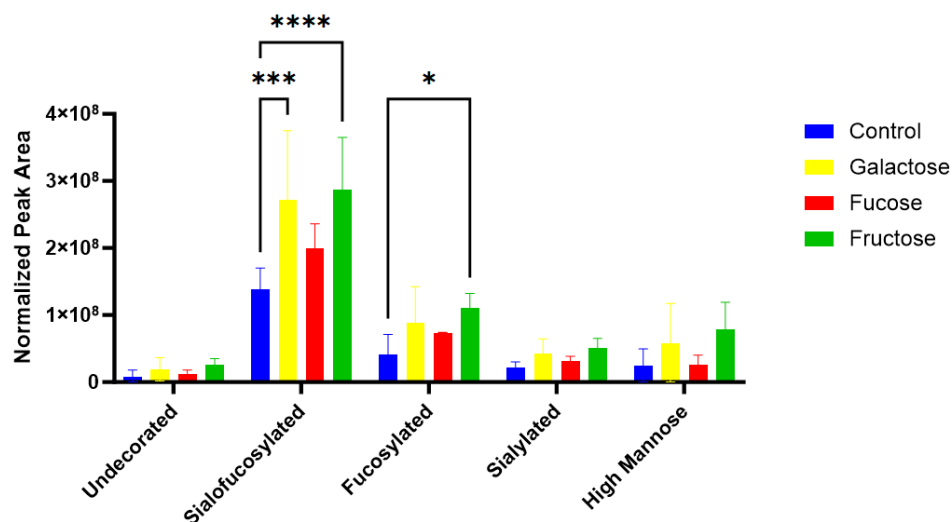


Figure 2.3. Bar graph representing the normalized peak areas or abundances of different glycan groups observed in LC-MS/MS analysis of cell surface glycans with different monosaccharide supplementations. The peak area integrated for every structure was normalized to the amounts of proteins it was extracted from so that it is not affected by changes in protein abundance, and it is based on the glycan abundance.

3.2. Glycoproteomics Analysis of the Secreted IgG1

Placing cells in serum free media allowed for extraction of antibodies that were secreted from the cell line. The purification process involved spin filtration and using magnetic protein A/G beads to extract IgG1 antibodies. The only observed glycopeptides from IgG1, were from the conserved site of Asn 297. N-Glycan at this site affect the biological functions of IgG1. Analysis of the control samples yielded the identification of 6 N-glycan structures at Asn 297,

and all of these structures were complex-type fucosylated N-glycans (Figure 2.5). Overall, the N-glycan structures observed on secreted IgG were not as diverse as the ones observed in the cell membrane. Additionally, the most prominent IgG1 N-glycan contained a bisecting GlcNAc in controls. Bisecting GlcNAc on Asn 297 N-glycans of IgG1 are considered to have pro-inflammatory effects.¹⁵

3.2.1. The Effects of Glycosylation Inhibitors on Secreted IgG1 Glycosylation

Glycoproteomics of IgG1 showed the changes in IgG1 N-glycans were similar to the ones observed in the cell membrane N-glycans with glycosylation inhibitors. This is the first report that shows the changes in glycosylation of cell membrane glycoproteins and secreted glycoproteins follow similar trends with metabolic engineering techniques. As the most abundant glycan structures observed at Asn 297 in control samples did not include sialic acid decorations, inhibition of sialylation did not affect the glycosylation at this site and the three most abundant structures at this site remained the same fucosylated structures (4510, 4410 and 3410), which were observed in the control (Figure 2.4). With the utilization of the fucosylation inhibitor, only afucosylated glycans were observed at this site, and the two most abundant structures at Asn 297 were 4500 and 4400, which are the afucosylated versions of the structures (4510 and 4410) observed in the control (Figure 2.4).

As the result of the kifunensine treatment, mostly high-mannose structures were observed at Asn 297 of IgG1, and Man-9 (9200) was the highest abundant glycan at this site with 92.2% relative abundance (Figure 2.4). Man-9 containing IgG1 is not commonly observed in human serum, and it may certainly exhibit unique structural and functional properties as it will be

presented in the modeling section below. As opposed to the cell membrane glycosylation, in which Man-5, Man-6, Man-7, and Man-8 (5200, 6200, 7200 and 8200) were among the most prominent glycans with kifunensine treatment (supplementary Figure 2.2), these high mannose glycans were not observed on IgG1, and the highest abundant glycans following Man-9 were 3410 and 4410 with 3% and 4% relative abundances respectively (Figure 2.4). 3410 and 4410 glycans were the second and third most prominent N-glycans on IgG1 in the control samples with 18% and 9% relative abundances respectively (Figure 2.4). This indicates the inhibitors are not affecting all glycoproteins in an identical manner, although overall changes in membrane and secreted IgG1 glycans are consistent in the context of the glycan types, the most abundant structures on IgG1 do not match with the ones observed in cell membrane glycomics after treatments.

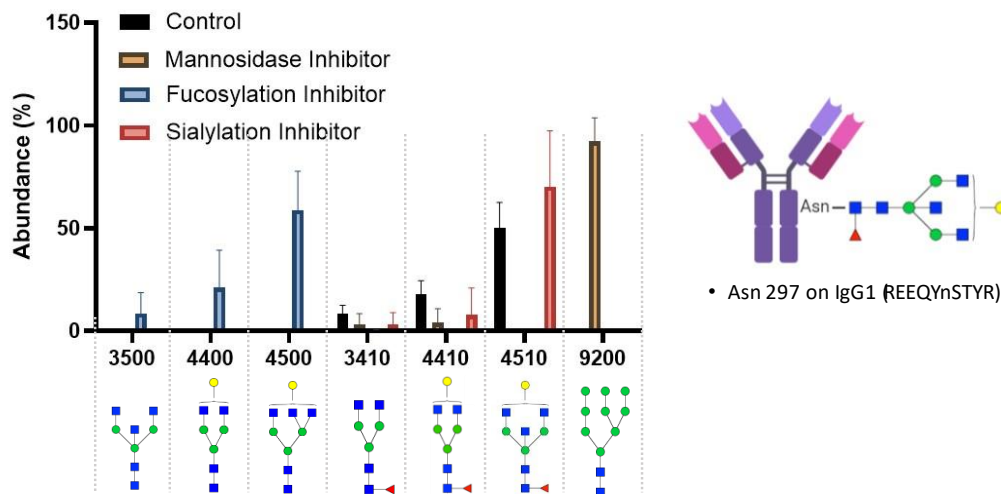


Figure 2.4. (Left) Bar graph representing the relative abundances of different glycans on Asn 297 of IgG using LC-MS/MS based glycoproteomics analysis of the secreted IgG1 from cells treated with different glycosylation inhibitors. The relative abundances were calculated based

on the precursor peak areas of every identified structure. The four-letter codes show the number of the building monosaccharides of the structures as follows HexHexNAcFucoseNeuAc. (Right) Representing most abundant glycan observed in control on the Asn (N) 297 of IgG1 with REEQYNSTYR tryptic digest sequence.

3.2.2. The Effects of Monosaccharide Supplementations on Secreted IgG1 Glycosylation

Monosaccharide supplementations also caused significant changes in IgG1 glycosylation. Although the cell surface glycomics results showed no significant changes with the fucose treatment, secreted IgG1 glycoproteomics showed all monosaccharide treatments (galactose, fructose and fucose) caused significant changes on glycosylation of IgG1. The changes included significant reductions in relative abundances of the most prominent structure (4510) with all three treatments, $p=0.0026$ for galactose and $p<0.0001$ for fructose and $p<0.0001$ for fucose, compared to the control (Figure 2.5). In addition to the significant decreases in the most prominent structure observed in the control, a new sialofucosylated N-glycan structure (5411) was observed in all monosaccharide treated samples, which was not observed in the control (Figure 2.5). Sialofucosylated IgG1 are known to have a higher half-life in serum compared to their un-sialyated counterparts.³⁰ Therefore, addition of monosaccharides to the cell stock for antibody production may certainly affect the biotherapeutic efficacy of the final product.

Glycosylation is heavily influenced by the abundances and activities of glycosylation-related enzymes, as well as the presence of the monosaccharide building blocks. These results show monosaccharides can affect abundances of glycan building blocks and cause changes in

glycosylation of glycoproteins. It appears the induced changes in N-glycans are glycoprotein specific, meaning the induced changes depend on the glycan structures that are 'default' or commonly built on a specific glycoprotein. An example of these structures is the fucosylated 4510 N-glycan structure on the secreted IgG1 from the CESS cells which makes up 72% of the N-glycan relative abundance in the controls (Figure 2.5)

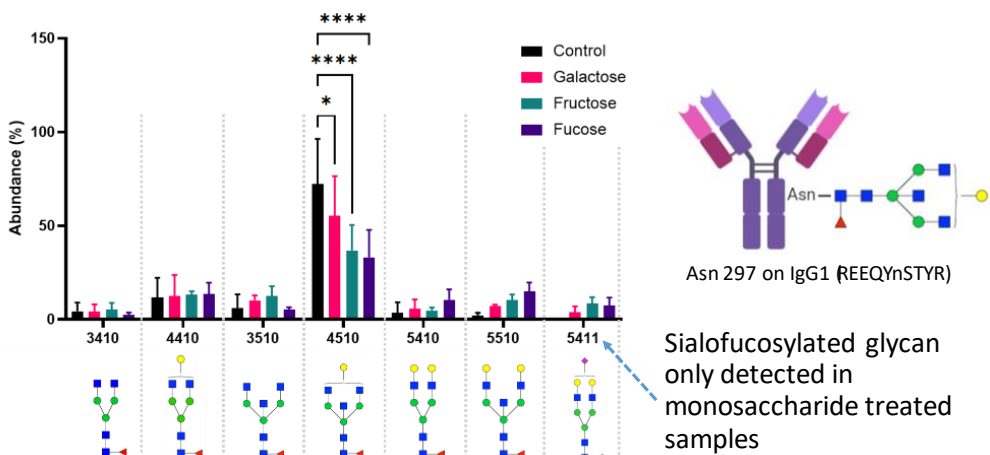
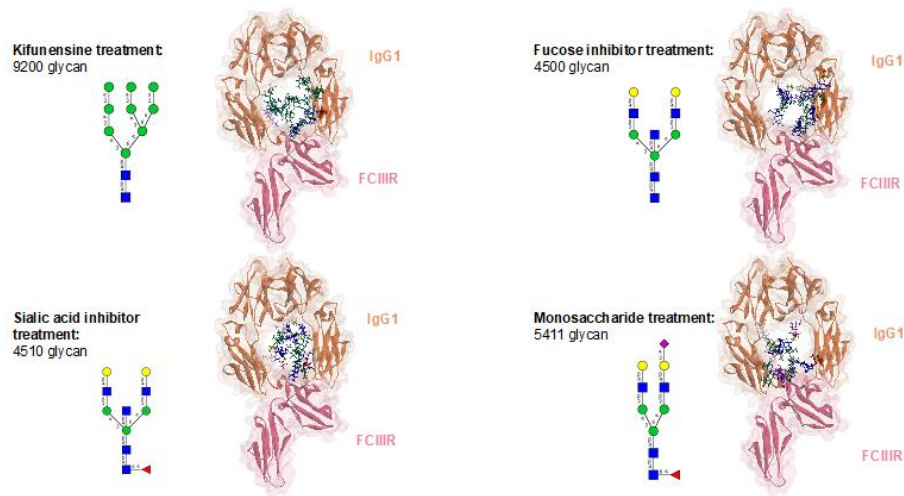


Figure 2.5. (Left) Bar graph representing the relative abundances of different glycans on Asn 297 of IgG using LC-MS/MS based glycoproteomics analysis of the secreted IgG1 from cells supplemented with different monosaccharides. The relative abundances were calculated based on the precursor peak areas of every identified structure. The four-letter codes show the number of the building monosaccharides of the structures as follows HexHexNAcFucoseNeuAc. (Right) Representing most abundant glycan observed in control on the Asn (N) 297 of IgG1 with REEQYNSTYR tryptic digest sequence.

3.3 Modeling IgG1 Fc interaction with IgG1 Fc Receptor

One of the major IgG1 functions is its binding to Fc binding receptors on the cell surface which leads to initiation of ADCC pathway and apoptosis. Studies have shown the binding affinity of IgG1 towards FcγRIIIb receptor is heavily affected by its glycosylation profile³¹. Absence of a single fucose residue has been shown to increase the binding affinity of IgG to the receptor by and increase ADCC by 50 folds.³² In this study, we observed a unique high mannose glycan structure (Man-9) which has not been reported previously in the literature. This glycan at Asn 297 of IgG may lead to exhibition of unique characteristics of the antibody. We modeled the binding of differentially glycosylated IgG1 to the FcγRIIIb receptor, with the most abundant structures observed in the control and sialic acid inhibitor treatment (4510), kifunensine treatment (9200), fucose inhibitor treatment (4500), and monosaccharide supplemented samples (5411) (Figure 2.6A). Our results show these three different glycans will certainly affect the IgG binding to Fc receptor differently. In particular, we observed the core-fucosylated structure 4510 in IgG1 contributed the most interactions with FcγRIIIb receptor, providing unique glycan-glycan interactions with the FcγRIIIb N-glycan (Figure 2.6B). Without the core-fucose in IgG1, the glycan-glycan interaction with FcγRIIIb is absent, potentially leading to reduced binding. Glycan-glycan interactions between IgG1 and FcγRIIIb have been reported to contribute to high binding affinity between IgG1 and FcγRIIIb receptor. Thus, directly modifying the IgG1 glycosylation with glyco-engineering provides an avenue for directly modulating functional activities of IgG1 antibodies.

A



B

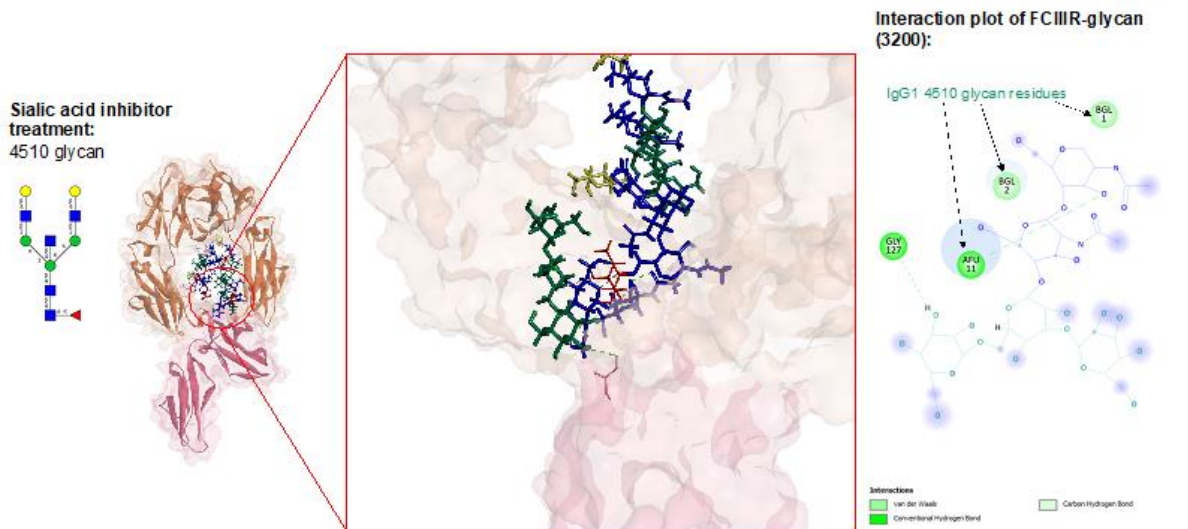


Figure 2.6. (A) The IgG1-FCIIIR complex were modeled with varying IgG1 glycosylation: control and sialylation inhibitor (4510), kifunensine treatment (9200), fucose inhibitor (4500), and monosaccharide supplemented (5411). (B) Glycan-glycan interactions between IgG1-4510 glycan and FCIIIR-3200 structures.

3.4 Transcriptomics Analysis of Glycosylation Inhibitor Treated Cells

Although changes in glycosylation are not expected to directly affect the transcriptomics of the cells, there has been no published study that has monitored the transcriptomics of cells through glycan changes induced by glycoenzyme inhibitors. Changes in glycosylation may potentially contribute to changes in the transcriptome. It is essential to monitor the viability of cells and changes in their macromolecular composition, as it has been shown certain glycan decorations, such as fucosylation, are essential for functions of certain cells with their inhibition being lethal *in vivo*.³³ The viability assay showed no significant changes in viability of cells in presence of the three inhibitors used in this experiment (supplementary Figure 2.6) signifying the CESS cell line can proliferate and operate with reduction in fucosylation, sialylation and undecoration of its N-glycan structures. The RNA data provided more intel into the induced changes in the cells. Based on the transcriptomics results, there were no significant changes ($P_{adj} < 0.05$) in abundances of any of the observed RNA with the FDR corrections to P values (Figure 2.7). This shows the induced changes in glycosylation did not have severe impacts on the transcriptomics of the cells overall. This aligns with our previous data where we monitored changes in protein expressions via proteomics of glycosylation inhibitor treated epithelial cells.⁹ Un-adjusted P values (without FDR corrections) showed 3-3.5% of transcriptomes were expressed significantly different in the glycosylation inhibited samples compared to the controls (Supplementary Figure 2.9). IGLC3 (immunoglobulin lambda constant 3) transcript expression was reduced significantly with the kifunensine treatment (Supplementary Figure 2.8). IGLC3 is a building block of the IgG1 light chain region. As mentioned previously, IgG1 with high mannose N-glycans are not generally observed in human body. The reduced expression of the IGLC3 RNA

with inhibition of glycan trimming enzymes may indicate signs of IgG1 mis-folding due to the absence of an untrimmed glycan at Asn 297, and the cells utilizing a feedback loop to reduce the expression of the glycoprotein due to misfolding.

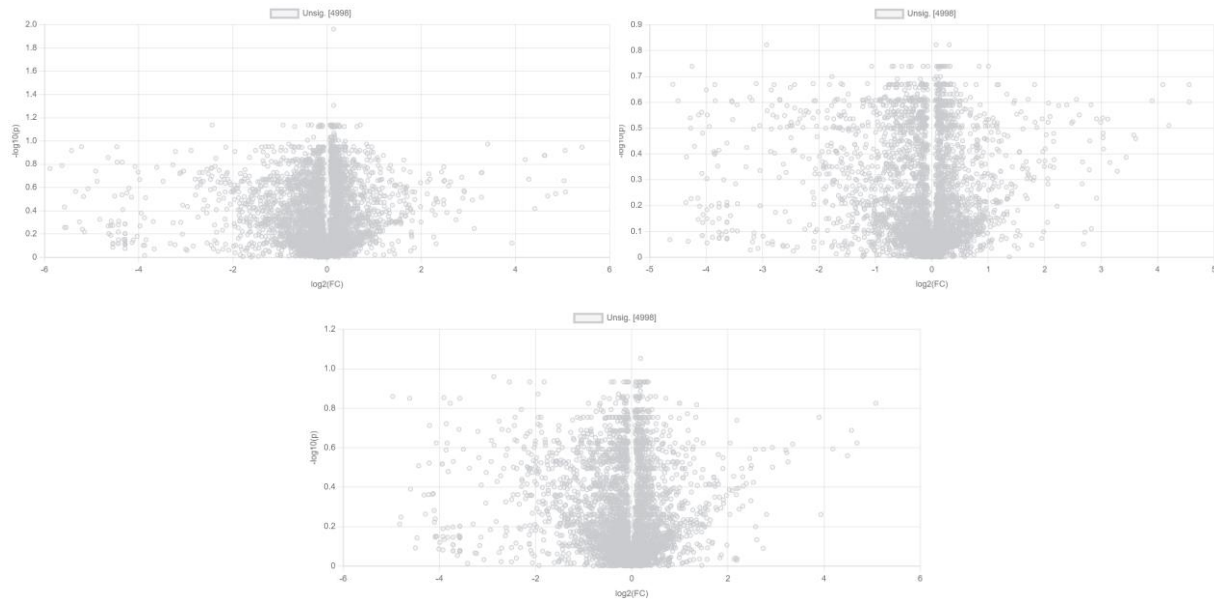


Figure 2.7. Volcano plots showing no drastic changes (fold change >2 and P_{adj} -value <0.05) in RNA expressions of cells after inhibitor treatments with FDR correction. Treatments of glycan trimming inhibitor (Top-Left), sialylation inhibitor (Top-Right) and fucosylation inhibitor (Bottom) are presented in the three plots with $\log_2(FC)$ as the X-axis and $-\log_{10}(P)$ as the Y-axis.

4. Discussion

The result of this study showed how glycosylation inhibitors alter both the cell surface glycans on the immune cells, and the glycan on the secreted IgG1. As expected, the inhibitors changed the glycan structures selectively by inhibiting a specific step in the glycosylation

pathway. With that said, the resulting structures on the cell membrane are not identical to the ones observed on the secreted antibody from the cell, signifying the glycoprotein specificity of glycosylation. A single glycoprotein, such as IgG1, may only contain a select group of glycans. The combination of the different glycans on different membrane glycoproteins leads to the heterogeneity and diversity of the N-glycan observed in the cell membrane.

The ability to glycoengineer both cell surface glycan structures and the secreted antibody glycans of the immune cells in parallel has significant therapeutic applications.²³ Immune cell and secreted antibody N-glycans undergo changes in diseases such as Asthma³⁴ and Rheumatoid arthritis,³⁵ and these changes often contribute to disease progression. The disease induced changes may potentially be reversed by the methods presented in this study, such as monosaccharide supplementations.

IgG1 is often utilized as a drug, targeting disease-inducing bacteria, proteins or even cancer cells, therefore being able to optimize IgG1 glycosylation is an essential step in developing its most efficient proteoform for optimal functionality. The types of glycans at Asn 297 of IgG affect its clearance rate in serum, for example IgG1 with sialylated N-glycans at Asn 297 have been shown to have a longer half-life than the ones with un-sialylated N-glycans at the same site,³⁶ therefore the glycan engineering methods demonstrated in this study can be used to produce therapeutic IgG1 antibodies with optimal clearance rate in serum.

Monitoring both the cell surface N-glycans and secreted antibody N-glycans of the immune cells in this study allowed for comparison of a single glycoprotein glycosylation with the overall cell membrane glycosylation. It is important to note glycosylation cannot be generalized for all glycoproteins and every glycoprotein may react differently to presence or absence of

monosaccharides substrates for glycosylations. For IgG1, presence of additional monosaccharides resulted in sialylation, indicating the cells favored incorporating sialic acids into the structures, perhaps due to increased structural stability of the glycoprotein with sialic acid incorporation. The same overall trend was observed with the cell surface glycans but with a broader number of N-glycan structures.

5. Conclusion

Glycosylation inhibitors induce selective changes in glycosylation of cell surface and secreted antibodies of immune cells. Changes in glycosylation did not cause drastic changes in RNA or viability of the cells. Additionally, a new Man-9 containing IgG1 was observed which exhibited unique structural characteristics based on the molecular dynamics modeling.

References:

- (1) Colley, K. J.; Varki, A.; Haltiwanger, R. S.; Kinoshita, T. Cellular Organization of Glycosylation. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.
- (2) Möckl, L. The Emerging Role of the Mammalian Glycocalyx in Functional Membrane Organization and Immune System Regulation. *Frontiers in Cell and Developmental Biology* **2020**, *8*.
- (3) Stanley, P.; Moremen, K. W.; Lewis, N. E.; Taniguchi, N.; Aebi, M. N-Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.
- (4) Toustou, C.; Walet-Balieu, M.; Kiefer-Meyer, M.; Houdou, M.; Lerouge, P.; Foulquier, F.; Bardor, M. Towards Understanding the Extensive Diversity of Protein N-glycan Structures in Eukaryotes. *Biol Rev Camb Philos Soc* **2022**, *97* (2), 732–748. <https://doi.org/10.1111/brv.12820>.
- (5) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. N-Glycans. In *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press, 1999.
- (6) Fisher, P.; Thomas-Oates, J.; Wood, A. J.; Ungar, D. The N-Glycosylation Processing Potential of the Mammalian Golgi Apparatus. *Frontiers in Cell and Developmental Biology* **2019**, *7*.
- (7) SOLÁ, R. J.; GRIEBENOW, K. Effects of Glycosylation on the Stability of Protein Pharmaceuticals. *J Pharm Sci* **2009**, *98* (4), 1223–1245. <https://doi.org/10.1002/jps.21504>.
- (8) Liu, F.-T.; Stowell, S. R. The Role of Galectins in Immunity and Infection. *Nat Rev Immunol* **2023**, 1–16. <https://doi.org/10.1038/s41577-022-00829-7>.
- (9) Zhou, Q.; Xie, Y.; Lam, M.; Lebrilla, C. B. N-Glycomic Analysis of the Cell Shows Specific Effects of Glycosyl Transferase Inhibitors. *Cells* **2021**, *10* (9), 2318. <https://doi.org/10.3390/cells10092318>.
- (10) Giron, L. B.; Tanes, C. E.; Schleimann, M. H.; Engen, P. A.; Mattei, L. M.; Anzurez, A.; Damra, M.; Zhang, H.; Bittinger, K.; Bushman, F.; Kossenkova, A.; Denton, P. W.; Tateno, H.;

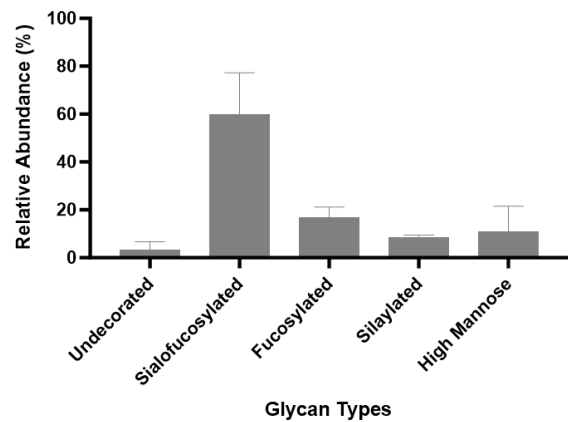
- Keshavarzian, A.; Landay, A. L.; Abdel-Mohsen, M. Sialylation and Fucosylation Modulate Inflammasome-Activating eIF2 Signaling and Microbial Translocation during HIV Infection. *Mucosal Immunol* **2020**, *13* (5), 753–766. <https://doi.org/10.1038/s41385-020-0279-5>.
- (11) Li, J.; Hsu, H.-C.; Ding, Y.; Li, H.; Wu, Q.; Yang, P.; Luo, B.; Rowse, A. L.; Spalding, D. M.; Bridges, S. L.; Mountz, J. D. Inhibition of Fucosylation Reshapes Inflammatory Macrophages and Suppresses Type II Collagen-Induced Arthritis. *Arthritis Rheumatol* **2014**, *66* (9), 2368–2379. <https://doi.org/10.1002/art.38711>.
- (12) *Sialylation - an overview | ScienceDirect Topics*.
<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/sialylation> (accessed 2024-02-06).
- (13) Büll, C.; den Brok, M. H.; Adema, G. J. Sweet Escape: Sialic Acids in Tumor Immune Evasion. *Biochim Biophys Acta* **2014**, *1846* (1), 238–246.
<https://doi.org/10.1016/j.bbcan.2014.07.005>.
- (14) Chia, S.; Tay, S. J.; Song, Z.; Yang, Y.; Walsh, I.; Pang, K. T. Enhancing Pharmacokinetic and Pharmacodynamic Properties of Recombinant Therapeutic Proteins by Manipulation of Sialic Acid Content. *Biomedicine & Pharmacotherapy* **2023**, *163*, 114757.
<https://doi.org/10.1016/j.biopha.2023.114757>.
- (15) Maverakis, E.; Kim, K.; Shimoda, M.; Gershwin, M. E.; Patel, F.; Wilken, R.; Raychaudhuri, S.; Ruhaak, L. R.; Lebrilla, C. B. Glycans in the Immune System and The Altered Glycan Theory of Autoimmunity: A Critical Review. *J Autoimmun* **2015**, *57*, 1–13.
<https://doi.org/10.1016/j.jaut.2014.12.002>.
- (16) Golay, J.; Andrea, A. E.; Cattaneo, I. Role of Fc Core Fucosylation in the Effector Function of IgG1 Antibodies. *Frontiers in Immunology* **2022**, *13*.
- (17) Pereira, N. A.; Chan, K. F.; Lin, P. C.; Song, Z. The “Less-Is-More” in Therapeutic Antibodies: Afucosylated Anti-Cancer Antibodies with Enhanced Antibody-Dependent Cellular Cytotoxicity. *MAbs* **2018**, *10* (5), 693–711.
<https://doi.org/10.1080/19420862.2018.1466767>.
- (18) Ruhaak, L. R.; Barkauskas, D. A.; Torres, J.; Cooke, C. L.; Wu, L. D.; Stroble, C.; Ozcan, S.; Williams, C. C.; Camorlinga, M.; Rocke, D. M.; Lebrilla, C. B.; Solnick, J. V. The Serum

- Immunoglobulin G Glycosylation Signature of Gastric Cancer. *EuPA Open Proteom* **2015**, *6*, 1–9. <https://doi.org/10.1016/j.euprot.2014.11.002>.
- (19) Alvarez, M. R. S.; Zhou, Q.; Tena, J.; Lebrilla, C. B.; Completo, G. C.; Heralde, F. M.; Cabanatan, M.; Barzaga, M. T.; Tan-Liu, N.; Ladrera, G. I.; Danguilan, J. L.; Rabajante, J.; Padolina, I.; Nacario, R. C. N-Glycan and Glycopeptide Serum Biomarkers in Philippine Lung Cancer Patients Identified Using Liquid Chromatography-Tandem Mass Spectrometry. *ACS Omega* **2022**, *7* (44), 40230–40240. <https://doi.org/10.1021/acsomega.2c05111>.
- (20) Tena, J.; Tang, X.; Zhou, Q.; Harvey, D.; Barajas-Mendoza, M.; Jin, L.; Maezawa, I.; Zivkovic, A. M.; Lebrilla, C. B. Glycosylation Alterations in Serum of Alzheimer’s Disease Patients Show Widespread Changes in N-glycosylation of Proteins Related to Immune Function, Inflammation, and Lipoprotein Metabolism. *Alzheimers Dement (Amst)* **2022**, *14* (1), e12309. <https://doi.org/10.1002/dad2.12309>.
- (21) Chakraborty, S.; Gonzalez, J.; Edwards, K.; Mallajosyula, V.; Buzzanco, A. S.; Sherwood, R.; Buffone, C.; Kathale, N.; Providenza, S.; Xie, M. M.; Andrews, J. R.; Blish, C. A.; Singh, U.; Dugan, H.; Wilson, P. C.; Pham, T. D.; Boyd, S. D.; Nadeau, K. C.; Pinsky, B. A.; Zhang, S.; Memoli, M. J.; Taubenberger, J. K.; Morales, T.; Schapiro, J. M.; Tan, G. S.; Jagannathan, P.; Wang, T. T. Proinflammatory IgG Fc Structures in Patients with Severe COVID-19. *Nat Immunol* **2021**, *22* (1), 67–73. <https://doi.org/10.1038/s41590-020-00828-7>.
- (22) Ash, M. K.; Bhimalli, P. P.; Cho, B.-K.; Mattamana, B. B.; Gambut, S.; Tarhoni, I.; Fhied, C. L.; Reyes, A. F.; Welninski, S. J.; Arivalagan, J.; Negrão, F.; Goel, R.; Beck, T. L.; Hope, T. J.; Sha, B. E.; Goo, Y. A.; Al-Harhi, L.; Mamede, J. I.; Borgia, J. A.; Kelleher, N. L.; Schneider, J. R. Bulk IgG Glycosylation Predicts COVID-19 Severity and Vaccine Antibody Response. *Cell Rep* **2022**, *41* (11), 111799. <https://doi.org/10.1016/j.celrep.2022.111799>.
- (23) Narimatsu, Y.; Büll, C.; Chen, Y.-H.; Wandall, H. H.; Yang, Z.; Clausen, H. Genetic Glycoengineering in Mammalian Cells. *J Biol Chem* **2021**, *296*, 100448. <https://doi.org/10.1016/j.jbc.2021.100448>.
- (24) Wong, M.; Xu, G.; Barboza, M.; Maezawa, I.; Jin, L.-W.; Zivkovic, A.; Lebrilla, C. B. Metabolic Flux Analysis of the Neural Cell Glycocalyx Reveals Differential Utilization of

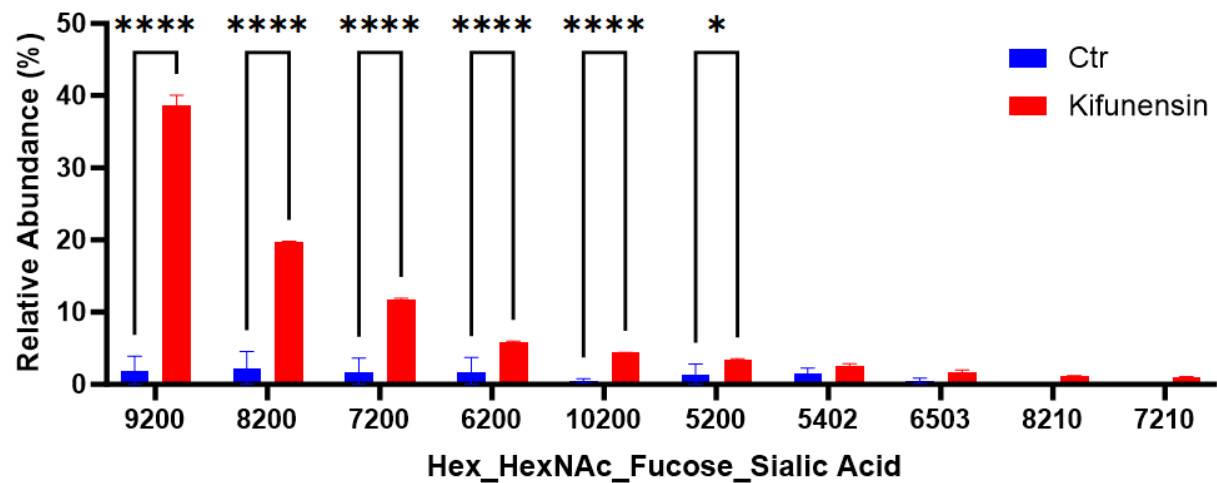
- Monosaccharides. *Glycobiology* **2020**, *30* (11), 859–871.
<https://doi.org/10.1093/glycob/cwaa038>.
- (25) Bagdonaite, I.; Malaker, S. A.; Polasky, D. A.; Riley, N. M.; Schjoldager, K.; Vakhrushev, S. Y.; Halim, A.; Aoki-Kinoshita, K. F.; Nesvizhskii, A. I.; Bertozzi, C. R.; Wandall, H. H.; Parker, B. L.; Thaysen-Andersen, M.; Scott, N. E. Glycoproteomics. *Nat Rev Methods Primers* **2022**, *2* (1), 1–29. <https://doi.org/10.1038/s43586-022-00128-4>.
- (26) Li, Q.; Xie, Y.; Wong, M.; Barboza, M.; Lebrilla, C. B. Comprehensive Structural Glycomic Characterization of the Glycocalyxes of Cells and Tissues. *Nat Protoc* **2020**, *15* (8), 2668–2704. <https://doi.org/10.1038/s41596-020-0350-4>.
- (27) Song, T.; Ozcan, S.; Becker, A.; Lebrilla, C. B. In-Depth Method for the Characterization of Glycosylation in Manufactured Recombinant Monoclonal Antibody Drugs. *Anal. Chem.* **2014**, *86* (12), 5661–5666. <https://doi.org/10.1021/ac501102t>.
- (28) Shah, N.; Kuntz, D. A.; Rose, D. R. Comparison of Kifunensine and 1-Deoxymannojirimycin Binding to Class I and II Alpha-Mannosidases Demonstrates Different Saccharide Distortions in Inverting and Retaining Catalytic Mechanisms. *Biochemistry* **2003**, *42* (47), 13812–13816. <https://doi.org/10.1021/bi034742r>.
- (29) Legler, K.; Rosprim, R.; Karius, T.; Eylmann, K.; Rossberg, M.; Wirtz, R. M.; Müller, V.; Witzel, I.; Schmalfeldt, B.; Milde-Langosch, K.; Oliveira-Ferrer, L. Reduced Mannosidase MAN1A1 Expression Leads to Aberrant N-Glycosylation and Impaired Survival in Breast Cancer. *Br J Cancer* **2018**, *118* (6), 847–856. <https://doi.org/10.1038/bjc.2017.472>.
- (30) Sheng, Y.; Vinjamuri, A.; Alvarez, M. R. S.; Xie, Y.; McGrath, M.; Chen, S.; Barboza, M.; Frieman, M.; Lebrilla, C. B. Host Cell Glycocalyx Remodeling Reveals SARS-CoV-2 Spike Protein Glycomic Binding Sites. *Front Mol Biosci* **2022**, *9*, 799703. <https://doi.org/10.3389/fmolb.2022.799703>.
- (31) Bas, M.; Terrier, A.; Jacque, E.; Dehenne, A.; Pochet-Béghin, V.; Béghin, C.; Dezetter, A.-S.; Dupont, G.; Engrand, A.; Beaufils, B.; Mondon, P.; Fournier, N.; Romeuf, C. de; Jorieux, S.; Fontayne, A.; Mars, L. T.; Monnet, C. Fc Sialylation Prolongs Serum Half-Life of Therapeutic Antibodies. *The Journal of Immunology* **2019**, *202* (5), 1582–1594. <https://doi.org/10.4049/jimmunol.1800896>.

- (32) Mizushima, T.; Yagi, H.; Takemoto, E.; Shibata-Koyama, M.; Isoda, Y.; Iida, S.; Masuda, K.; Satoh, M.; Kato, K. Structural Basis for Improved Efficacy of Therapeutic Antibodies on Defucosylation of Their Fc Glycans. *Genes Cells* **2011**, *16* (11), 1071–1080. <https://doi.org/10.1111/j.1365-2443.2011.01552.x>.
- (33) Shields, R. L.; Lai, J.; Keck, R.; O’Connell, L. Y.; Hong, K.; Meng, Y. G.; Weikert, S. H. A.; Presta, L. G. Lack of Fucose on Human IgG1 N-Linked Oligosaccharide Improves Binding to Human FcγRIII and Antibody-Dependent Cellular Toxicity. *J Biol Chem* **2002**, *277* (30), 26733–26740. <https://doi.org/10.1074/jbc.M202069200>.
- (34) Schneider, M.; Al-Shareffi, E.; Haltiwanger, R. S. Biological Functions of Fucose in Mammals. *Glycobiology* **2017**, *27* (7), 601–618. <https://doi.org/10.1093/glycob/cwx034>.
- (35) Dammen-Brower, K.; Epler, P.; Zhu, S.; Bernstein, Z. J.; Stabach, P. R.; Braddock, D. T.; Spangler, J. B.; Yarema, K. J. Strategies for Glycoengineering Therapeutic Proteins. *Front Chem* **2022**, *10*, 863118. <https://doi.org/10.3389/fchem.2022.863118>.
- (36) Xie, X.; Kong, S.; Cao, W. Targeting Protein Glycosylation to Regulate Inflammation in the Respiratory Tract: Novel Diagnostic and Therapeutic Candidates for Chronic Respiratory Diseases. *Frontiers in Immunology* **2023**, *14*.
- (37) Kissel, T.; Toes, R. E. M.; Huizinga, T. W. J.; Wuhler, M. Glycobiology of Rheumatic Diseases. *Nat Rev Rheumatol* **2023**, *19* (1), 28–43. <https://doi.org/10.1038/s41584-022-00867-4>.
- (38) Liu, Y. D.; Flynn, G. C. Effect of High Mannose Glycan Pairing on IgG Antibody Clearance. *Biologicals* **2016**, *44* (3), 163–169. <https://doi.org/10.1016/j.biologicals.2016.02.003>.

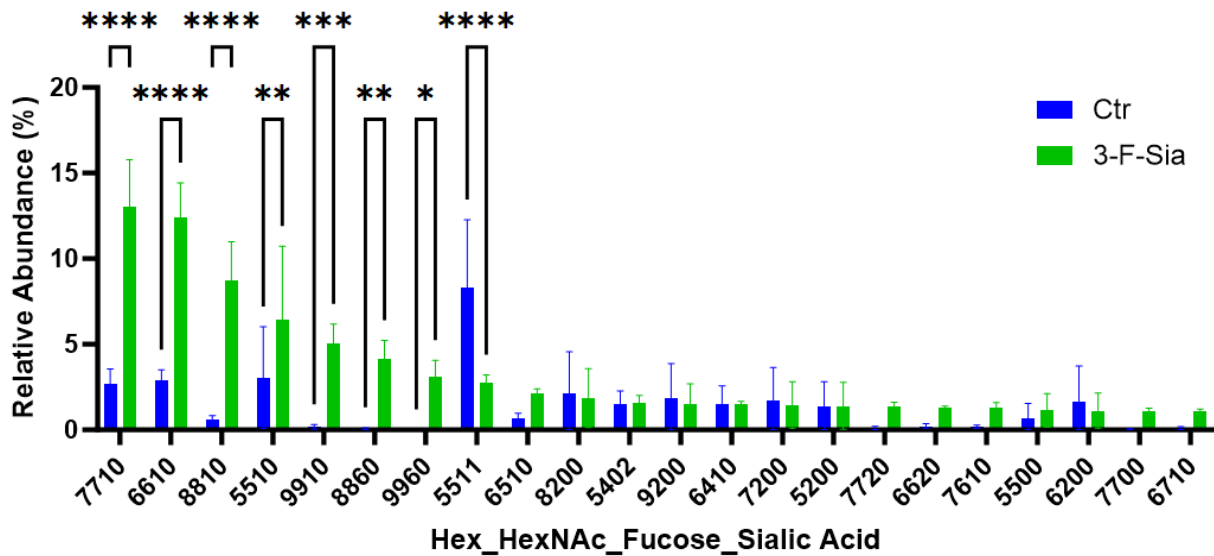
Supplementary Figures



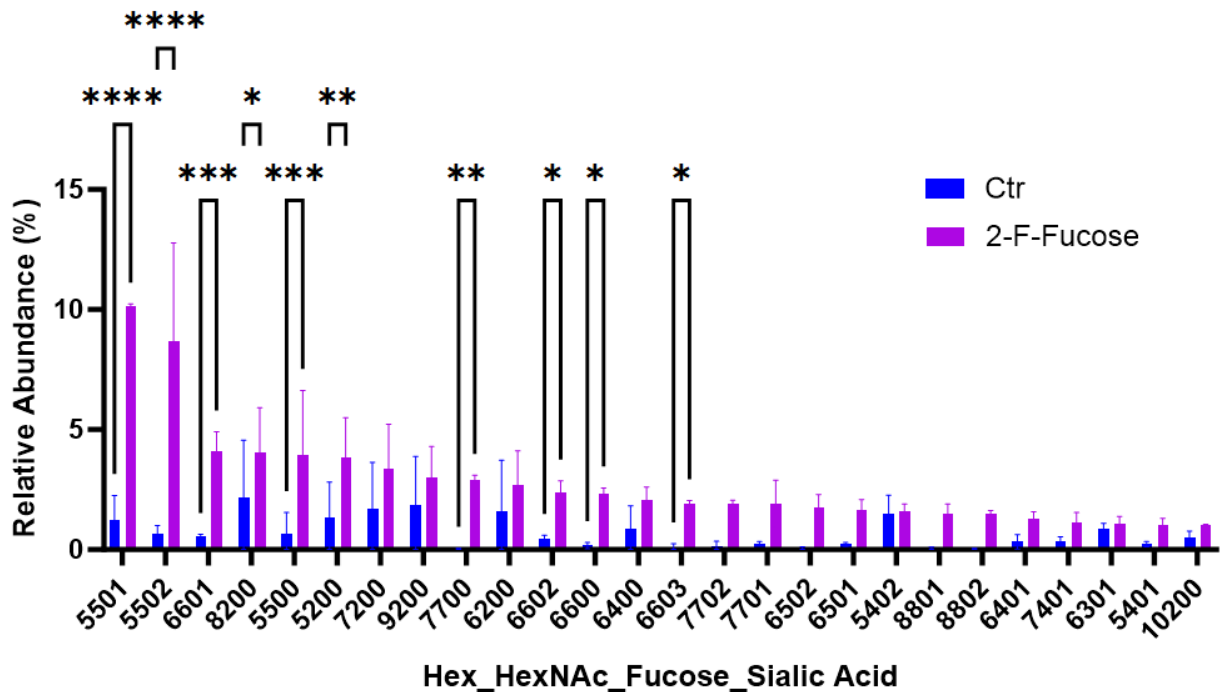
Supplementary Figure 2.1. Relative abundances of N-glycan types observed in CESS cell membrane.



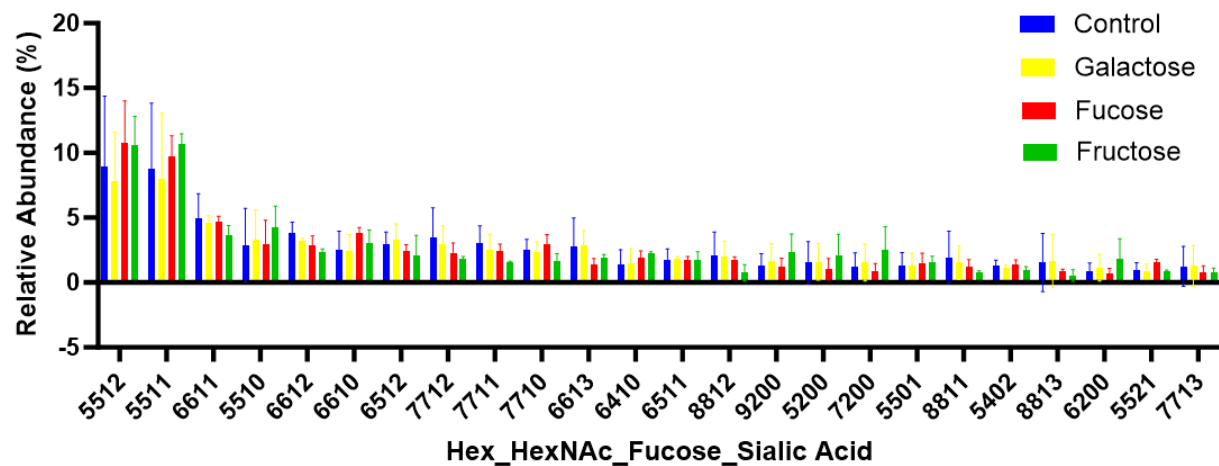
Supplementary Figure 2.2. Relative abundances of membrane N-glycan compositions with greater than 1% abundances in kifunensine treated cells.



Supplementary Figure 2.3. Relative abundances of membrane N-glycan compositions with greater than 1% abundances in sialylation inhibitor treated cells.



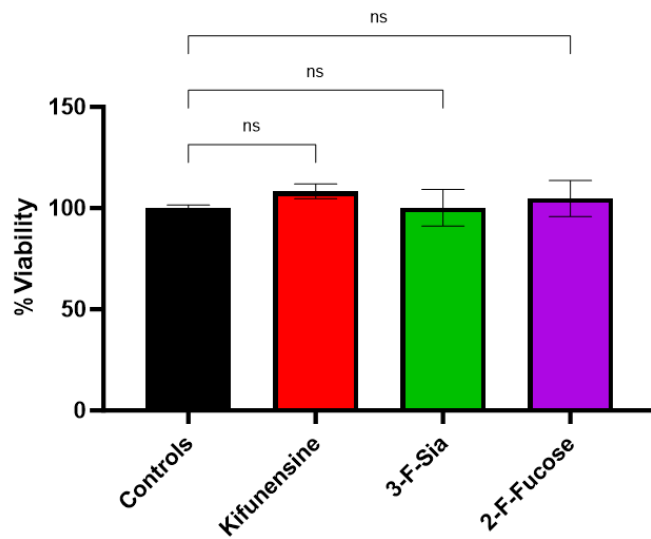
Supplementary Figure 2.4. Relative abundances of membrane N-glycan compositions with greater than 1% abundances in fucosylation inhibitor treated cells.



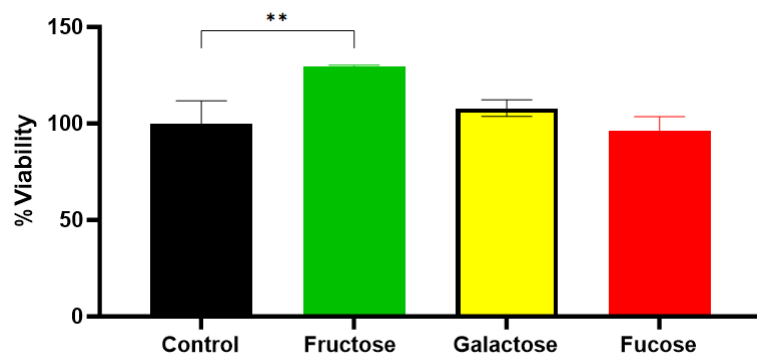
Supplementary Figure 2.5. Relative abundances of membrane N-glycan compositions with greater than 1% abundances in Control, compared to monosaccharide treated cells.

Supplementary Table 2.1. N-Glycan with significant changes in their relative abundances due to monosaccharide supplementations. The four-letter codes show the number of the building monosaccharides of the structures as follows HexHexNAcFucoseNeuAc. FC represents the fold change (abundances in treatment over abundances in control). P values are determined based on the T-test.

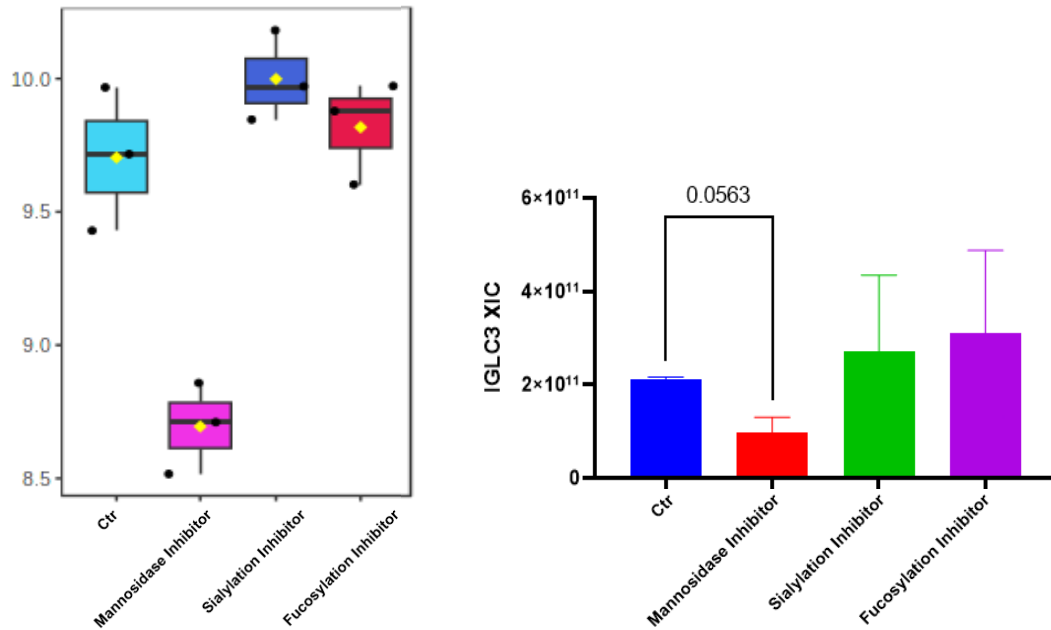
Treatment	Glycan	FC	P _{raw}	Glycan Type
Galactose	6623	2.4009	0.013396	SF
	7412	6.7336	0.014448	SF
	7601	3.4641	0.049415	S
Fucose	7502	6.1709	0.0012194	S
	6620	4.4128	0.0029998	F
	9920	5.9904	0.0032585	F
	8821	6.1755	0.023964	SF
	11210	4.316	0.027219	F
	6310	2.3793	0.032038	F
	6521	3.2286	0.033454	SF
	7621	7.0076	0.036994	SF
	8820	14.213	0.04812	F
	6641	2.4125	0.048786	F
Fructose	5522	2.0815	0.020192	SF
	6520	3.2386	0.027485	F
	6622	0.53355	0.037836	SF
	5310	2.7998	0.040122	F
	6612	0.61104	0.042859	SF
	9813	14.64	0.047719	SF



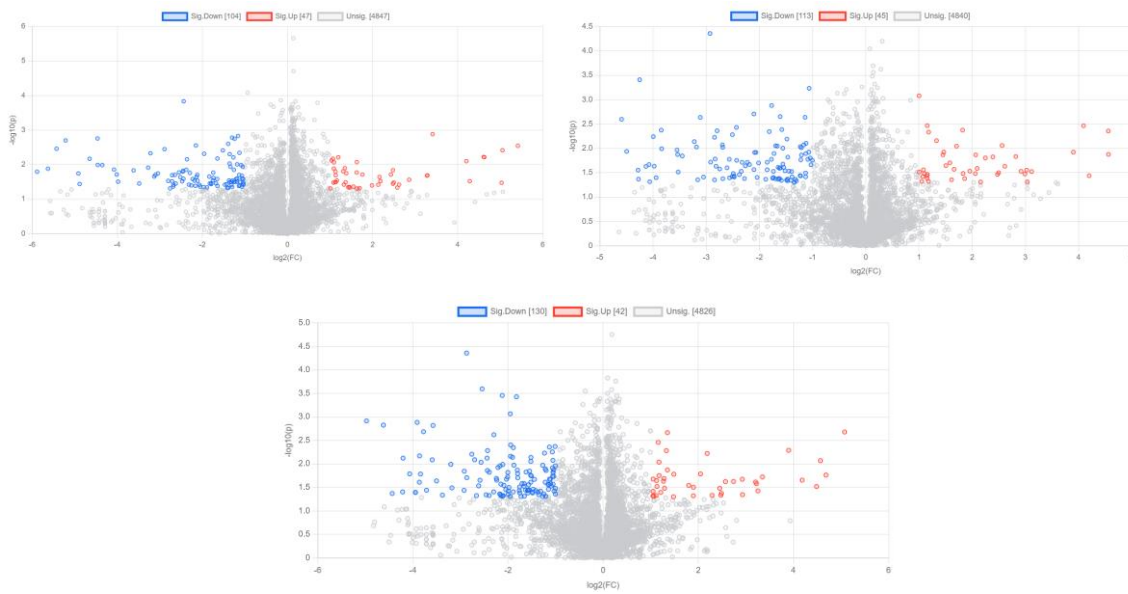
Supplementary Figure 2.6. Cell viability of inhibitor treated cells showing no significant change in cell viability with treatments.



Supplementary Figure 2.7. Cell viability of monosaccharide supplemented cells showing a significant increase in cell viability with fructose treatment.



Supplementary Figure 2.8. (Left) Normalized RNA expression of IgG1 building block (IGLC3) abundance in inhibitor treated samples. (Right) Peak areas of identified IGLC3 peptides in glycoproteomic analysis of extracted IgG1 after inhibitor treatments showing a potential reduction ($p=0.056$) in IGLC3 production with kifunensine/mannosidase inhibitor.



Supplementary Figure 9. Volcano plots showing changes (fold change >2 and P-value <0.05) in RNA expressions of cells after inhibitor treatments without FDR corrections. Treatments of kifunensine (Left), sialylation inhibitor (Right) and fucosylation inhibitor (Bottom) are presented in the three plots with Log₂(FC) as the X-axis and -Log₁₀(P) as the Y-axis.

Chapter III

Pemphigus-associated desmoglein-specific IgG1 and IgG4 have a dominant agalactosylated glycan modification

Authors: Armin Oloumi^{1*}, Stephanie T. Le^{2*}, Yiyun Liu¹, Samantha Herbert², Antonio Ji-Xu², Alexander A. Merleev², Olga Kruglinskaya³, Kimberly Artounian², Dayoung Park^{1,4}, Alina I. Marusina², John J. Zone⁵, Ron Feldman⁶, Carlito B. Lebrilla¹, Emanuel Maverakis²

1. Department of Chemistry, University of California, Davis, CA
2. Department of Dermatology, University of California, Davis School of Medicine, Sacramento, CA
3. Physioseq, Sacramento, CA
4. Department of Surgery, Harvard Medical School, Boston, MA
5. Department of Dermatology, University of Utah, Salt Lake City, UT
6. Department of Dermatology, Emory University School of Medicine, Atlanta, GA

*Authors contributed equally

Abstract:

Aberrant glycosylation of antibodies is observed in various diseases, such as Rheumatoid Arthritis. Identifying glycosylation changes in diseases allows for development of glycan biomarkers. By characterizing these glycan changes, contribution of glycans to disease progression can also be further understood. This is especially important in auto-immune diseases because the glycans on immunoglobulins affect their functions. In this study, we identified and quantified total immunoglobulin glycosylation changes in control and Pemphigus Vulgaris patients (PV, n=10) sera, then further analyzed the glycosylation of disease specific antibodies using ELISA and LC-MS/MS methods. The total immunoglobulin glycosylation analysis revealed the upregulation of an agalactosylated glycan structure (3410) on all classes of IgG in PV. Similarly, the glycoproteome analysis of disease specific IgG1 and IgG4 validated the presence of the same glycan on both antibodies. Furthermore, the abundance of a sialylated N-glycan (4411) on IgG1 negatively correlated with skin-specific autoantibody titers, thought to be a surrogate for disease activity.

Introduction:

Aberrant glycan modifications have been hypothesized to drive autoimmune pathogenesis, especially those at the conserved immunoglobulin constant heavy chain domain 2 (CH2)-84.4 site of IgG.¹ Glycans at this site dictate Ig effector function, with agalactosylated glycans (lacking galactose residues) being linked to antibody-mediated autoimmune conditions such as rheumatoid arthritis.² Pemphigus vulgaris is an autoimmune blistering disease characterized by autoantibodies directed against desmosomal cadherins. The glycosylation of these disease specific antibodies is not well understood.

We have previously demonstrated the use of dynamic multiple reaction monitoring LC-MS/MS methods for characterization of site-specific changes in glycosylation of immunoglobulins.³⁻⁵ These methods have been used to identify glycan changes in sera of gastric and lung cancer,^{6,7} as well as Alzheimer's disease patients.⁸ Given the importance of glycans on antibody functions, it is crucial to characterize the glycan changes in auto immune disease. Identifying the glycan changes unravels their contributions to disease pathology and provides addition insights for diagnosis and therapeutic developments.

Although the use of total immunoglobulin glycosylation characterization methods (dMRM) has been demonstrated previously, identifying glycosylation of disease specific antibodies is more challenging. IgG glycoproteins have a conserved N-glycan site in the Fc region, therefore tryptic digests of total and auto-reactive IgG yield indistinguishable glycopeptides that have the same amino acid sequence. To characterize the auto-reactive antibody glycans, auto-reactive antibodies need to be isolated from the rest of serum antibodies first. ELISA plates

coated with disease target proteins, such as DSG-3 in pemphigus vulgaris, are used to isolate disease specific antibodies.

To identify Ig glycan alterations in pemphigus, we obtained de-identified serum samples from patients with pemphigus vulgaris and healthy controls. DSG-specific autoantibody titers were determined by standardized ELISA assay. Skin-directed autoantibody titers were determined by indirect immunofluorescence (IIF), using monkey esophagus and normal skin as substrate. Glycoproteomic profiling was performed by Multiple Reaction Monitoring (MRM) mass spectrometry (MS) and orbitrap MS.^{3–5,9,10} To our knowledge, this is the first study which identifies both total immunoglobulin glycosylation changes and auto-reactive antibody glycosylation changes of a disease. Having both sets of glycosylation data allows us to confirm the glycan biomarker observed in total serum correspond to the one observed in autoantibodies of PV.

Materials and Methods:

Serum sample preparation

Serum sample preparation procedure was adapted from a previous publication.³ Briefly, 100 μ L of 50 mM NH_4HCO_3 (Millipore Sigma, cat. no. 09830) was added to 2 μ L of each serum sample prior to digestion. Serum proteins were reduced using 2 μ L of 550 mM DTT (Promega, cat. no. V3151) in a 60°C water bath for 50 min and alkylated using 4 μ L of 450 mM IAA (MilliporeSigma, cat. no. 144-48-9) at room temperature for 30 min in the dark. Afterwards, 1 μ g of trypsin (Promega, cat. no. V5111) in 2 μ L of 50 mM NH_4HCO_3 was added and proteins were

digested in a 37 °C water bath for 18 h. For absolute quantification, 100 µg of IgG, IgA and IgM (Sigma-Aldrich, St. Louis, MO) was digested according to the same described protocol and a dilution series was made prior to sample injection.

Indirect Immunofluorescence (IIF)

Patient serum was progressively diluted beginning at 1:5 in four two-fold screening dilutions, layered on sections of intact human skin, and monkey esophagus substrates and reacted with fluorescein isothiocyanate (FITC) - conjugated antibodies to IgG and IgA. When positive, the serum is further diluted in two-fold reductions to the limiting dilution of antibody detection or to a maximum dilution of 1:40,960. The limiting- dilution, end point titer is reported for each substrate.

Desmoglein ELISA

Desmoglein 1 and desmoglein 3 antibody levels determined by FDA-approved enzyme linked immunosorbent assays (ELISA) (DSG 1 & DSG 3 ELISA Test System, MBL BION) according to the manufacturer's instructions.

UPLC-ESI-QqQ-MS analysis of Pemphigus serum samples

The MRM MS method used for this study requires predetermined knowledge of the peptide or glycopeptide's LC retention time and its collision induced dissociation (CID) behavior, which we have previously determined for all the non-glycosylated peptides and glycopeptides used in this study.³⁻⁵ The specific method used herein has been highly validated and the

monitored transitions have been described in detail previously.^{1,4,5} The list of monitored transitions are presented in Supplemental Table S3.

The neat enzymatically prepared samples containing both peptides and glycopeptides were then directly analyzed without further hands-on sample cleanup or dilution using an Agilent 1290 infinity liquid chromatography (LC) system coupled to an Agilent 6490 triple quadrupole (QqQ) mass spectrometer (Agilent Technologies, Santa Clara, CA), as previously described.^{3–5,11} Briefly, an Agilent Eclipse plus C18 (RRHD 1.8 μ m, 2.1 $\times$ 100 mm) coupled with an Agilent Eclipse plus C18 pre-column (RRHD 1.8 μ m, 2.1 $\times$ 5 mm) was used for UPLC separation. 1.0 μ L of the digested serum samples was injected and analyzed using a 25-min binary gradient consisting of solvent A of 3% acetonitrile, 0.1% formic acid, solvent B of 90% acetonitrile, 0.1% formic acid in nano-pure water (v/v) at a flow rate of 0.5 mL/min.

Results were integrated using Agilent MassHunter Quantitative Analysis B.5.0 software. Protein concentrations were determined based on calibration curves and glycopeptide relative responses were calculated using the area under the curves of the glycopeptide and a non-glycosylated reference peptide from the same protein.

Enrichment of DSG 3-specific Antibodies

DSG 3-specific antibodies were enriched using DSG 3-coated ELISA plates (MBL Bion, IL, USA). Pooled sera of the patient (n=11) and control group (n=6) were prepared. Each pooled sera sample (5 μ L) were diluted to 100 μ L using the sample diluent solution from the ELISA kit and applied to the ELISA plate wells. After a one-hour incubation at room temperature, the wells were washed 4 times (300 μ L each) with the wash buffer solution from the ELISA kit, 3 times (300 μ L

each) with DPBS (GIBCO, MA, USA) and 3 times (300 μ L each) with 50 mM ammonium bicarbonate solution.

Digestion of DSG 3-bound proteins

In-plate digestion was performed according to Li et al.¹² with slight modifications. Briefly, 30 μ L of 50 mM ammonium bicarbonate solution was added to each well. The proteins were reduced with 2 μ L of 550 mM dithiothreitol (DTT) and incubated at 55° C for 50 min, then they were alkylated with 4 μ L of 450 mM iodoacetamide (IAA) and incubated at room temperature for 20 min. The samples were digested with 2 μ g of trypsin (Promega, MI, USA) and incubated at 37 °C for 18 hours. The digested samples were dried using vacuum centrifugation.

Glycoproteomic analysis with Orbitrap LC-MS/MS

The method for untargeted glycoproteomics analysis was adapted from Li et al.¹² Briefly, the peptide and 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, United States). The analytes were separated on an Acclaim™ PepMap™ 100 C18 LC Column (3 μ m, 0.075 mm $\times$ 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 1.8 kV spray voltage, 275°C ion transfer capillary temperature, and the acquisition was performed with the full MS scanned from 700 to 2000 m/z 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.

Analysis of Orbitrap LC-MS/MS data

Glycopeptide fragmentation spectra were annotated using Byonic software (Protein Metrics, CA, United States) against the protein sequences. Common modifications, including cysteine carbamidomethyl, methionine oxidation, asparagine deamidation and glutamine deamidation were assigned. A published in-house N-glycan library¹² was utilized for the glycopeptide identification, and relative abundance values were calculated based on the glycopeptide precursor peak areas.

Statistical analysis

Datasets related to this article can be found at <https://data.mendeley.com/datasets/cvyhwh6d65/2>, hosted at Mendeley Data. All statistical analyses were done using R software.¹³ For each glycovariant, a robustly quantified non-glycosylated peptide was used as an internal reference for calculating each glycovariant's relative abundance. Trypsin-digested protein standards were used to calculate each protein's absolute abundance. For each analyte, skewedness was calculated, and data was log transformed when necessary to remove excessive skewness. Outliers were identified using "extremevalues" R package,¹⁴ and when present, were winsorized from the analysis, so that the outliers were set equal to the nearest non-outlier value. Analytes could be detected in all samples, thus, there was no need for imputation of missing data. ANCOVA and linear regression assumptions about the normality of residuals were examined by use of the Shapiro–Wilk test. Collinearity of variables in

the multivariate models was examined by calculating variance inflation factor (excessive if > 2.5) with “car” R package.¹⁵

Data sharing

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD04.

Results:

1. Glycoproteomics of antibodies in pemphigus vulgaris patients using MRM LC-MS/MS

MRM MS (Figure 3.1a) was used to monitor the abundances of 52 glycopeptides and 10 non-glycosylated peptides (Supplemental Figure S1) in serum samples obtained from pemphigus vulgaris patients (n=10, Supplemental Table 3.1) and age- and sex-matched healthy controls (n=30, Supplemental Figure 3.2). To identify disease-specific glycan alterations, multiple linear regression was performed, with glycopeptides and non-glycosylated peptides as dependent variables in the models. After correcting for multiple comparisons, seven glycopeptides were found to be differentially expressed (Supplemental Table 3.2). Of these, the agalactosylated IgG CH2-84.4-modifying glycan g:3410 was strongly upregulated in pemphigus across all IgG subclasses, IgG1, IgG2, and IgG3/4 ($q = 7.7\text{e-}11$, $6.7\text{e-}10$, $1.0\text{e-}05$, respectively; Supplemental Table 3.2 and Figure 3.1b). In fact, the singleton classifier, IgG1 g:3410, was highly accurate in discriminating pemphigus vulgaris patients from healthy controls (area under respective Receiver Operating Characteristic (ROC) curve = 1, Figure 3.1c).

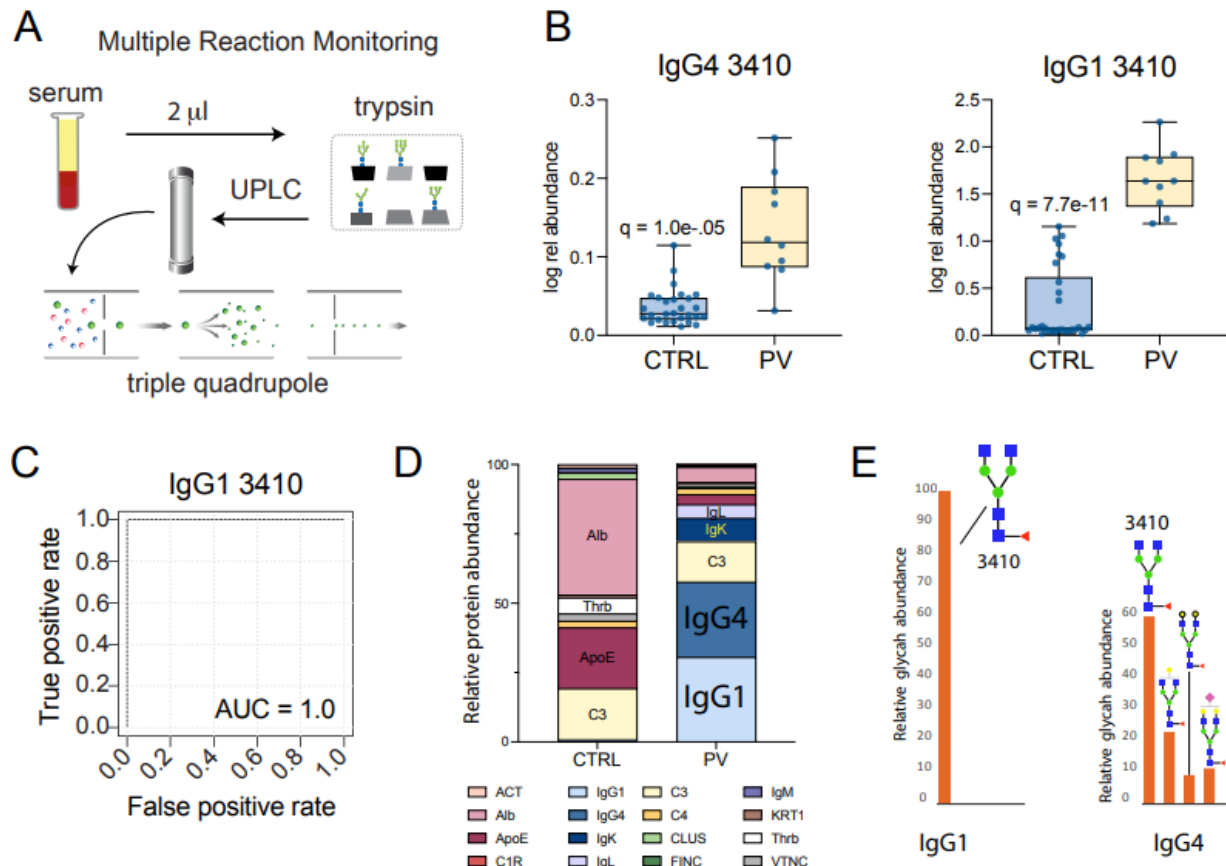


Figure 3.1. Multiple Reaction Monitoring (MRM) mass spectrometry identifies differentially expressed glycovariants in pemphigus vulgaris. (a) MRM method. **(b)** The relative abundance of IgG-modifying glycans g:3410 is increased in pemphigus. The Y-axis represents log relative abundance of the glycopeptide. Glycans modifications are presented as a 4-digit number where the number of hexose, N-acetylglucosamine, fucose, and sialic acid residues is indicated sequentially. **(c)** Receiver operating characteristic curve demonstrating the ability of IgG1 g:3410 to distinguish individuals with pemphigus vulgaris from healthy controls. **(d)** Orbitrap MS profiling of trypsin-treated DSG3-coated wells incubated with either pooled healthy control or pooled pemphigus vulgaris serum. DEG3 and trypsin detected fragments removed prior to analysis. **(e)** glycoprofiling of DSG3-specific IgG1 and IgG4 autoantibodies detected in “d”. CTRL,

control; PV, pemphigus vulgaris; q = false discovery rate (FDR) corrected p value; AUC, area under the curve.

2. Comparison of glycosylation changes in antibodies of pemphigus vulgaris (PV) and primary biliary cirrhosis (PBC) patients

To determine if these results were specific to pemphigus, results were compared to primary biliary cirrhosis (PBC), a disease characterized by autoreactive antibodies directed against mitochondria (Supplementary Figure 3.3). Both PBC and pemphigus had increases in the relative abundance of IgG1 g:3510 and decreases in the relative abundances of IgM 209 g:5511 and IgM 209 g:5512. However, IgG1 g:3410 was not significantly elevated in PBC and IgG4 g:3410 was downregulated in PBC, the latter being opposite to that observed in pemphigus. Finally, IgM and J Chain were increased in PBC, but not in pemphigus (Supplementary Figure 3.3).

3. Glycoproteomics of disease specific autoantibodies in pemphigus vulgaris patients using ELISA and LC-MS/MS

To determine if the g:3410 glycan modification was relevant to pemphigus pathophysiology, an additional independently acquired pemphigus serum set (n=11) was collected for glycoprofiling of autoreactive DSG-specific IgG [University of Utah samples (whole serum glycoprofiling), Emory University samples (DSG-specific Ig glycoprofiling)]. DSG3-coated plates were then incubated with either pooled pemphigus serum or healthy control serum. After washing, the remaining bound proteins were treated with trypsin and the resulting mixture of

peptides and glycopeptides was analyzed by orbitrap MS. Immunoglobulin fragments were not detected after DSG plates were incubated with control healthy serum, indicating that healthy controls do not have DSG-specific autoreactive Ig. In contrast, abundant IgG1 and IgG4 peptides, at roughly a 1:1 ratio, were detected after DSG3-coated plates were incubated with pooled pemphigus serum. Thus, across a population of pemphigus patients, DSG-specific IgG is predominantly of the IgG1 and IgG4 subclasses (Figure 3.1d). Consistent with the targeted MRM glycoprofiling of pemphigus serum, orbitrap MS glycoprofiling of DSG-specific Ig revealed that 100% of DSG3-specific IgG1 and the majority of DSG3-specific IgG4 was modified by g:3410 (Figure 3.1e).

4. Correlation of sialylated Ig glycans with autoantibody titers in pemphigus vulgaris

Pemphigus disease severity has been reported to correlate with autoantibody titers.^{16,17} In the current dataset, both anti-DSG titers (determined by ELISA) and anti-monkey esophagus titers (determined by IIF) positively correlated with human skin-specific autoantibody IIF titers (Supplemental Figure 3.4), which demonstrates that these putative measures of disease activity had strong concordance. We then sought to determine if the relative abundance of any of the differentially expressed Ig glycopeptides correlated with autoantibody titers to normal skin. This revealed a strong negative correlation between skin-specific IgG and the relative abundance of sialylated IgG1 g:4411 and IgG3/4 g:4411 ($r_s = -0.89$, false discovery rate (FDR) = $8.5e-3$; $r_s = -0.76$, $p = 1.2e-2$, respectively Figure 3.2a).

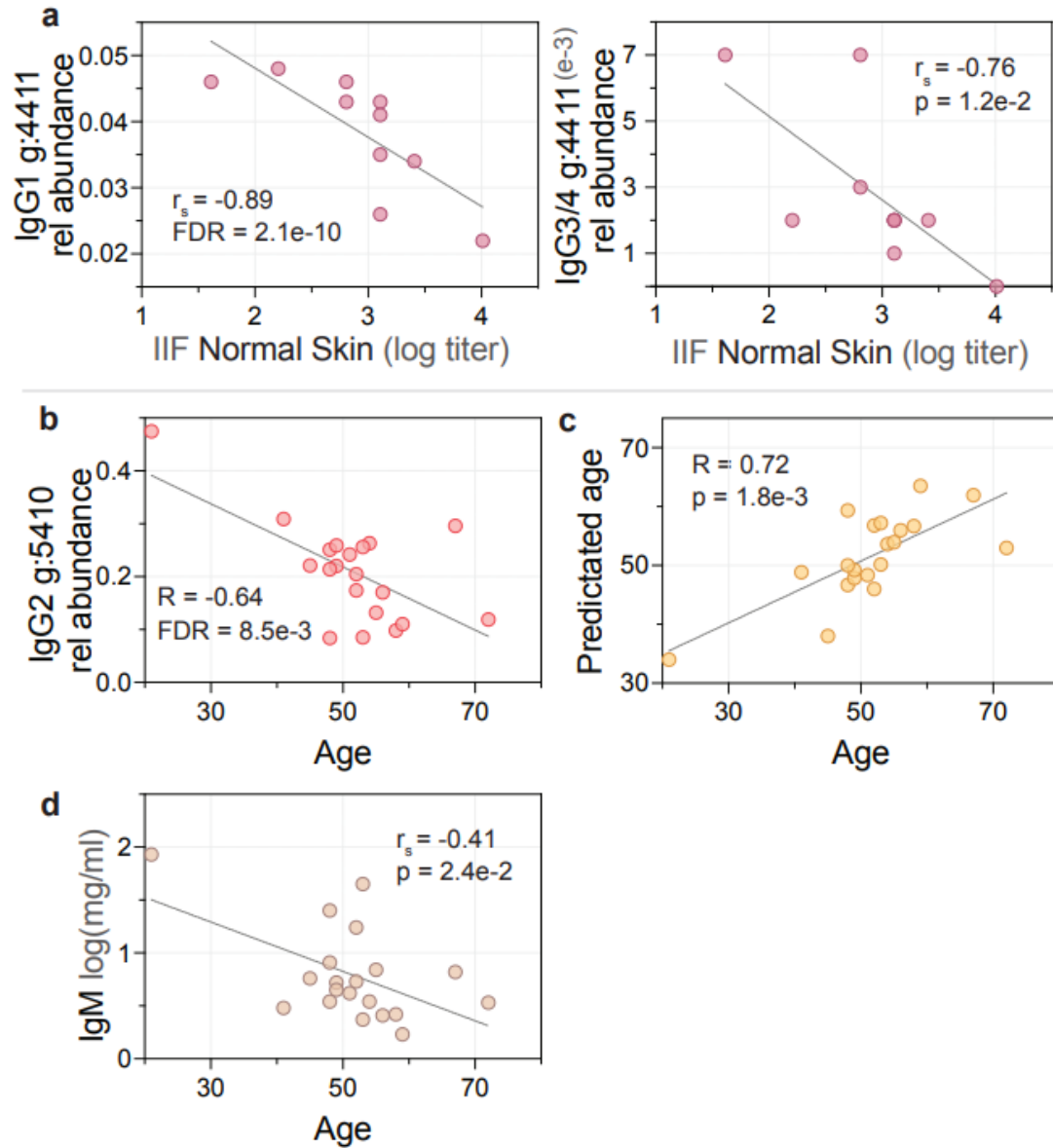


Figure 3.2. Specific sialylated Ig glycan modifications correlate with autoantibody titers in pemphigus vulgaris. (a) The relative abundance of sialylated IgG1 g:4411 and IgG3/4 g:4411 negatively correlate with normal skin-specific autoantibody titers. **(b)** Multiple logistic regression including glycopeptide, age, gender, and disease state identifies a negative association between IgG2 g:5410 and age. **(c)** Two-analyte age prediction regression model constructed from previously identified age-associated glycopeptides. **(d)** Scatter plot demonstrates a negative

relationship between age and serum IgM. IIF, indirect immunofluorescence antibody titers; esoph, esophagus; DSG1, antibody titers against desmoglein-1 detected by ELISA; Ig, immunoglobulin; FDR, false discovery rate.

In addition to disease-specific glycan alterations, linear regression analysis also identified age-associated glycopeptides (e.g. IgG2 G:5410, Figure 3.2b). An age prediction model ($\text{Age} = 42(\text{IgG2g:3510}) - 30(\text{IgG1g:5410}) + 60$) also performed well ($r = 0.72$, $p = 1.8\text{e-}3$) (Figure 3.2c). Further interrogation of the model revealed low multicollinearity among its variables (average variance inflation factor [VIF] = 1.11) and diagnostic plots of the model revealed good linearity, a lack of overly influential cases, and homoscedastic data, respectively (Supplementary Figure 3.5). Also, similar to previous studies,^{9,18} a negative association between serum IgM and age was observed ($r_s = -0.41$, $p = 2.4\text{e-}2$, Figure 3.2e).

Discussion

Our findings can be summarized as follows: 1) in pemphigus, the agalactosylated CH2-84.4 g:3410 N-glycan modification was increased across all IgG subclasses; 2) IgG1 g:3410 titers can identify pemphigus vulgaris patients from healthy controls; 3) the g:3410 glycan modification is present at the CH2-84.4 glycosylation of virtually all DSG3-specific IgG1; 4) the abundance of sialylated IgG1 g:4411 negatively correlates with skin-specific autoantibody titers, thought to be a surrogate for disease activity; 5) only a subset of pemphigus-associated glycan alterations can be detected in PBC; and 6) similar to previous reports, there is a well-defined set of Ig glycans that are associated with age. It is highly likely that the agalactosylated g:3410 glycan modification

of DSG3-specific IgG1 contributes to its pathogenicity, as Ig CH2-84.4 glycan modifications terminating in N-acetylglucosamine residues facilitate Ig-mediated complement activation and Ig Fc receptor binding¹. Also, the elevation of agalactosylated g:3410 across all IgG (i.e. DSG-specific and non-skin-specific IgG) maybe one reason why patients with pemphigus are predisposed to other antibody-mediated autoimmune diseases. Specifically, sialylated glycan modifications at CH2-84.4 bestow IgG with anti-inflammatory properties,¹ which is likely why sialylated IgG1 g:4411 negatively correlates with skin-specific autoantibody titers, a surrogate for disease severity.

Conclusion:

The utility of the glycan alterations described herein as predictive biomarkers can be explored in future studies, e.g. to determine their ability to identify patients with a low likelihood of disease recurrence following discontinuation of immunosuppressive therapy. Finally, our findings linking glycan alterations to skin-specific autoantibodies raises the possibility of targeting glycosylation pathways to treat pemphigus vulgaris and other antibody-mediated autoimmune diseases. The lack of pemphigus disease severity scores and other clinical data metrics are limitations of our current study, preventing us from firmly establishing a direct link between Ig glycan titers and disease severity.

REFERENCES

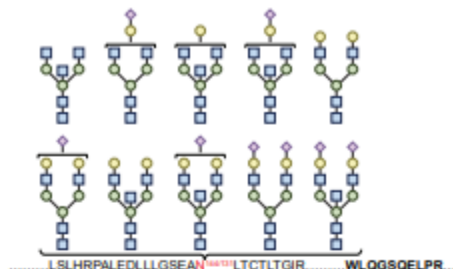
- (1) Maverakis, E.; Kim, K.; Shimoda, M.; Gershwin, M. E.; Patel, F.; Wilken, R.; Raychaudhuri, S.; Ruhaak, L. R.; Lebrilla, C. B. Glycans in the Immune System and The Altered Glycan Theory of Autoimmunity: A Critical Review. *Journal of Autoimmunity* **2015**, *57*, 1–13. <https://doi.org/10.1016/j.jaut.2014.12.002>.
- (2) Gudelj, I.; Salo, P. P.; Trbojević-Akmačić, I.; Albers, M.; Primorac, D.; Perola, M.; Lauc, G. Low Galactosylation of IgG Associates with Higher Risk for Future Diagnosis of Rheumatoid Arthritis during 10 years of Follow-Up. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2018**, *1864* (6, Part A), 2034–2039. <https://doi.org/10.1016/j.bbadis.2018.03.018>.
- (3) Hong, Q.; Lebrilla, C. B.; Miyamoto, S.; Ruhaak, L. R. Absolute Quantitation of Immunoglobulin G and Its Glycoforms Using Multiple Reaction Monitoring. *Anal Chem* **2013**, *85* (18), 8585–8593. <https://doi.org/10.1021/ac4009995>.
- (4) Hong, Q.; Ruhaak, L. R.; Stroble, C.; Parker, E.; Huang, J.; Maverakis, E.; Lebrilla, C. B. A Method for Comprehensive Glycosite-Mapping and Direct Quantitation of Serum Glycoproteins. *J Proteome Res* **2015**, *14* (12), 5179–5192. <https://doi.org/10.1021/acs.jproteome.5b00756>.
- (5) Li, Q.; Kailemia, M. J.; Merleev, A. A.; Xu, G.; Serie, D.; Danan, L. M.; Haj, F. G.; Maverakis, E.; Lebrilla, C. B. Site-Specific Glycosylation Quantitation of 50 Serum Glycoproteins Enhanced by Predictive Glycopeptidomics for Improved Disease Biomarker Discovery. *Anal. Chem.* **2019**, *91* (8), 5433–5445. <https://doi.org/10.1021/acs.analchem.9b00776>.
- (6) Ruhaak, L. R.; Barkauskas, D. A.; Torres, J.; Cooke, C. L.; Wu, L. D.; Stroble, C.; Ozcan, S.; Williams, C. C.; Camorlinga, M.; Rocke, D. M.; Lebrilla, C. B.; Solnick, J. V. The Serum Immunoglobulin G Glycosylation Signature of Gastric Cancer. *EuPA Open Proteom* **2015**, *6*, 1–9. <https://doi.org/10.1016/j.euprot.2014.11.002>.
- (7) Alvarez, M. R. S.; Zhou, Q.; Tena, J.; Lebrilla, C. B.; Completo, G. C.; Heralde, F. M. I.; Cabanatan, M.; Barzaga, Ma. T.; Tan-Liu, N.; Ladrera, G. I.; Danguilan, J. L.; Rabajante, J.; Padolina, I.; Nacario, R. C. N-Glycan and Glycopeptide Serum Biomarkers in Philippine Lung

- Cancer Patients Identified Using Liquid Chromatography–Tandem Mass Spectrometry. *ACS Omega* **2022**, 7 (44), 40230–40240. <https://doi.org/10.1021/acsomega.2c05111>.
- (8) Tena, J.; Tang, X.; Zhou, Q.; Harvey, D.; Barajas-Mendoza, M.; Jin, L.-W.; Maezawa, I.; Zivkovic, A. M.; Lebrilla, C. B. Glycosylation Alterations in Serum of Alzheimer’s Disease Patients Show Widespread Changes in N-Glycosylation of Proteins Related to Immune Function, Inflammation, and Lipoprotein Metabolism. *Alzheimers Dement (Amst)* **2022**, 14 (1), e12309. <https://doi.org/10.1002/dad2.12309>.
 - (9) Merleev, A. A.; Park, D.; Xie, Y.; Kailemia, M. J.; Xu, G.; Ruhaak, L. R.; Kim, K.; Hong, Q.; Li, Q.; Patel, F.; Wan, Y.-J. Y.; Marusina, A. I.; Adamopoulos, I. E.; Lal, N. N.; Mitra, A.; Le, S. T.; Shimoda, M.; Luxardi, G.; Lebrilla, C. B.; Maverakis, E. A Site-Specific Map of the Human Plasma Glycome and Its Age and Gender-Associated Alterations. *Scientific Reports* **2020**, 10 (1), 17505. <https://doi.org/10.1038/s41598-020-73588-x>.
 - (10) Park, D. D.; Xu, G.; Wong, M.; Phoomak, C.; Liu, M.; Haigh, N. E.; Wongkham, S.; Yang, P.; Maverakis, E.; Lebrilla, C. B. Membrane Glycomics Reveal Heterogeneity and Quantitative Distribution of Cell Surface Sialylation. *Chem. Sci.* **2018**, 9 (29), 6271–6285. <https://doi.org/10.1039/C8SC01875H>.
 - (11) Miyamoto, S.; Stroble, C. D.; Taylor, S.; Hong, Q.; Lebrilla, C. B.; Leiserowitz, G. S.; Kim, K.; Ruhaak, L. R. Multiple Reaction Monitoring for the Quantitation of Serum Protein Glycosylation Profiles: Application to Ovarian Cancer. *J. Proteome Res.* **2018**, 17 (1), 222–233. <https://doi.org/10.1021/acs.jproteome.7b00541>.
 - (12) Li, Q.; Xie, Y.; Wong, M.; Lebrilla, C. B. Characterization of Cell Glycocalyx with Mass Spectrometry Methods. *Cells* **2019**, 8 (8), 882. <https://doi.org/10.3390/cells8080882>.
 - (13) R Foundation for Statistical Computing V, Austria. . R Development Core Team (2008) R: A Language and Environment for Statistical Computing., [Http://Www.R-Project.Org.](Http://Www.R-Project.Org;); 2008.
 - (14) Van Der Loo M. Extremevalues, an R Package for Outlier Detection in Univariate Data. . Version 2.3 Ed. <Http://Www.Github.Com/Markvanderloo/Extremevalues>.2010.
 - (15) Fox J, Weisberg S. An R Companion to Applied Regression: Sage Publications, 2018.

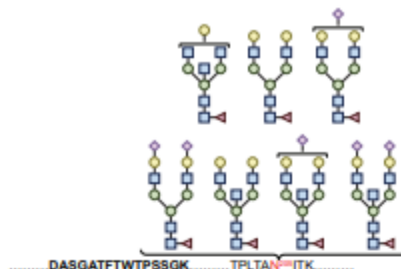
- (16) Aksu, D.; Peksari, Y.; Arica, I. E.; Gurgey, E. ASSESSING THE AUTOANTIBODY LEVELS IN RELATION TO DISEASE SEVERITY AND THERAPY RESPONSE IN PEMPHIGUS PATIENTS. *Indian J Dermatol* **2010**, 55 (4), 342–347. <https://doi.org/10.4103/0019-5154.74536>.
- (17) Weiss, D.; Ristl, R.; Griss, J.; Bangert, C.; Foedinger, D.; Stingl, G.; Brunner, P. M. Autoantibody Levels and Clinical Disease Severity in Patients with Pemphigus: Comparison of Aggregated Anti-Desmoglein ELISA Values and Indirect Immunofluorescence Titres. *Acta Derm Venereol* **2015**, 95 (5), 559–564. <https://doi.org/10.2340/00015555-2023>.
- (18) Listì, F.; Candore, G.; Modica, M. A.; Russo, M.; Di Lorenzo, G.; Esposito-Pellitteri, M.; Colonna-Romano, G.; Aquino, A.; Bulati, M.; Lio, D.; Franceschi, C.; Caruso, C. A Study of Serum Immunoglobulin Levels in Elderly Persons That Provides New Insights into B Cell Immunosenescence. *Ann N Y Acad Sci* **2006**, 1089, 487–495. <https://doi.org/10.1196/annals.1386.013>.

Supplementary Figures

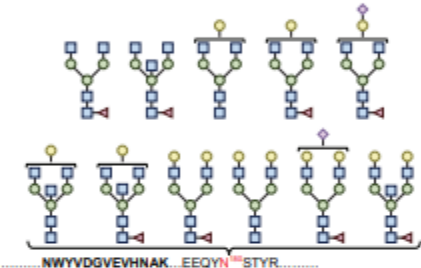
Immunoglobulin heavy constant alpha 1/2 (IgA1/2)



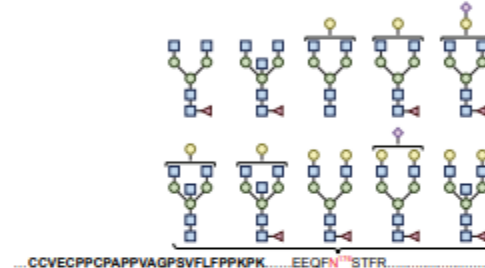
Immunoglobulin heavy constant alpha 2 (IgA2)



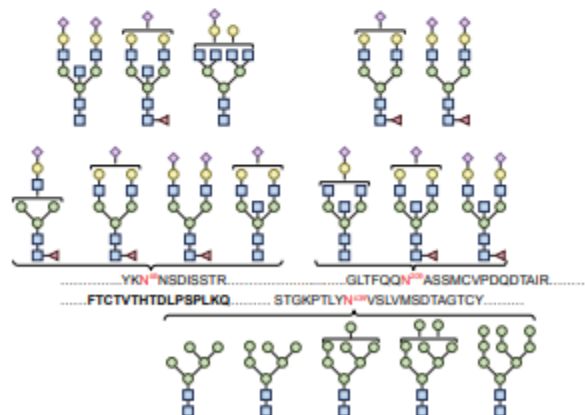
Immunoglobulin heavy constant gamma 1 (IgG1)



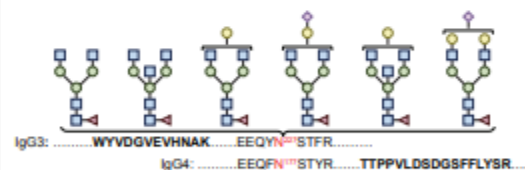
Immunoglobulin heavy constant gamma 2 (IgG2)



Immunoglobulin heavy constant mu (IgM)



Immunoglobulin heavy constant gamma 3/4 (IgG3/4)



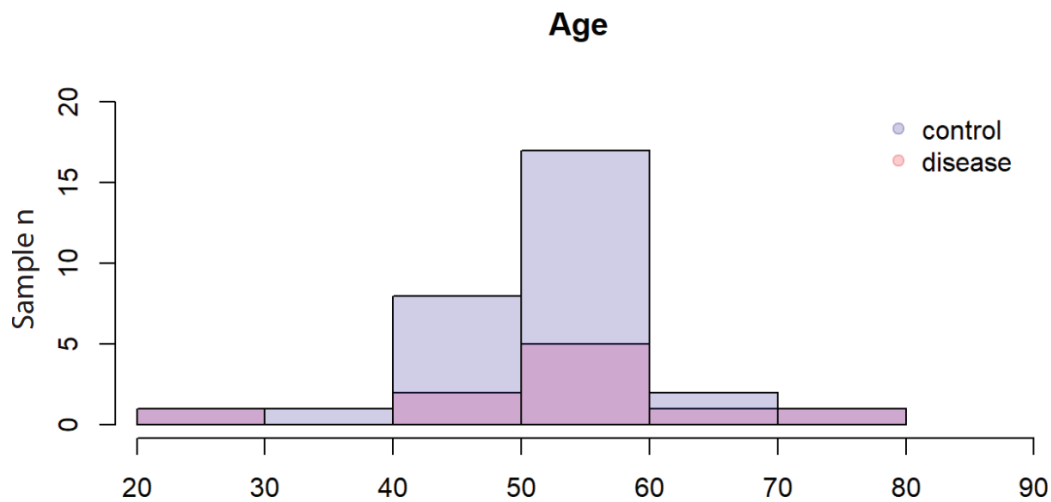
Immunoglobulin J chain



Figure Legend

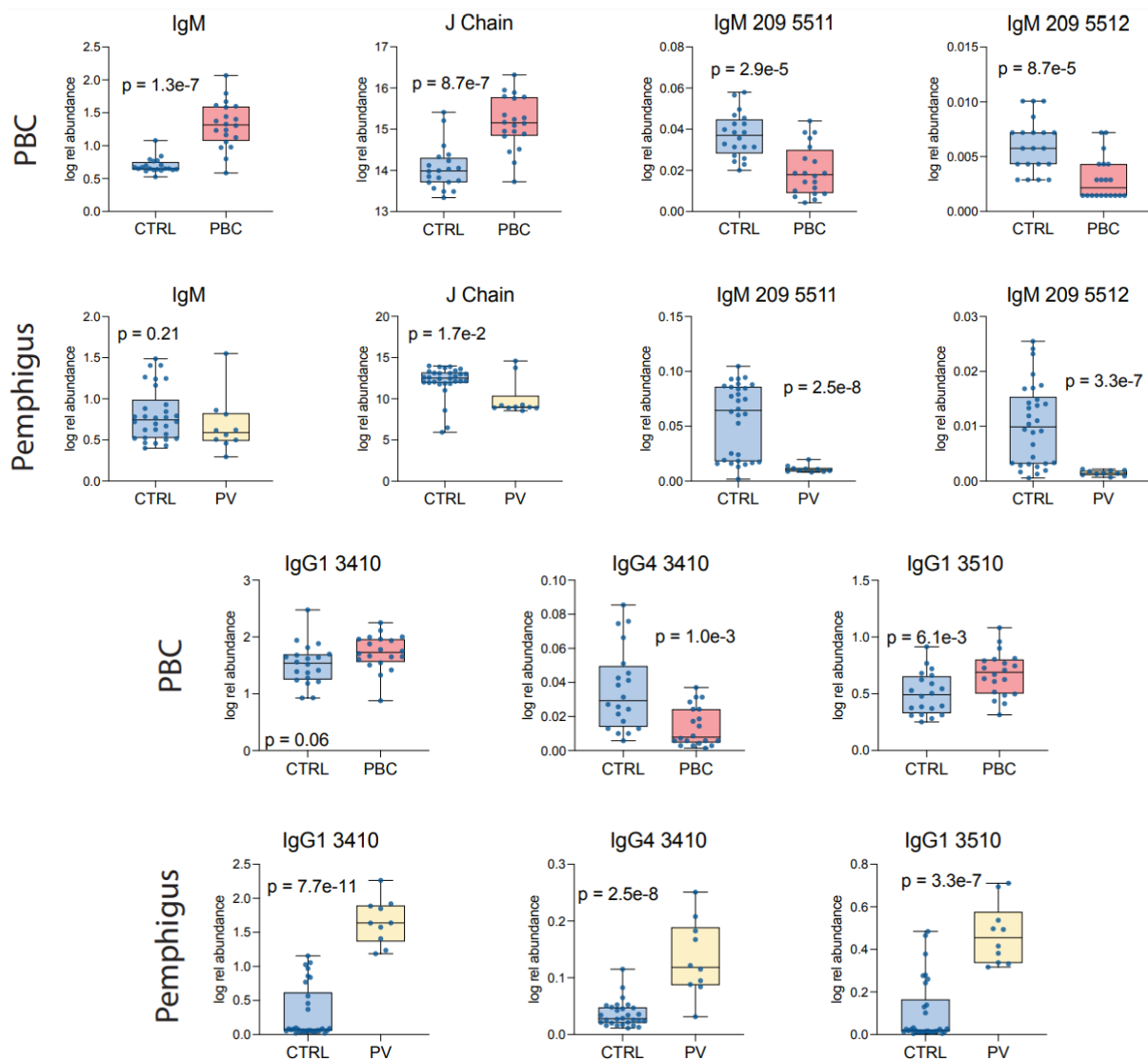
- | | | |
|---------------------|-----------|-------------------------|
| N-Acetylglucosamine | Galactose | N-Acetylneuraminic acid |
| Mannose | Fucose | N-Acetylgalactosamine |

Supplemental Figure 3.1. Structures and peptide sequences of site-specific glycans monitored, including all IgG subclasses, IgA, IgM and J chain. IgG3 and IgG4 subclasses are combined because the amino acid sequence of their trypsin-generated CH2 84.4 glycopeptide are identical. Likewise, IgA1 and IgA2 are also combined. The bolded sequences represent non-glycosylated reference peptides used to calculate the relative abundance of each glycopeptide. These non-glycosylated reference peptides are constant across all Ig subclasses of the same class.

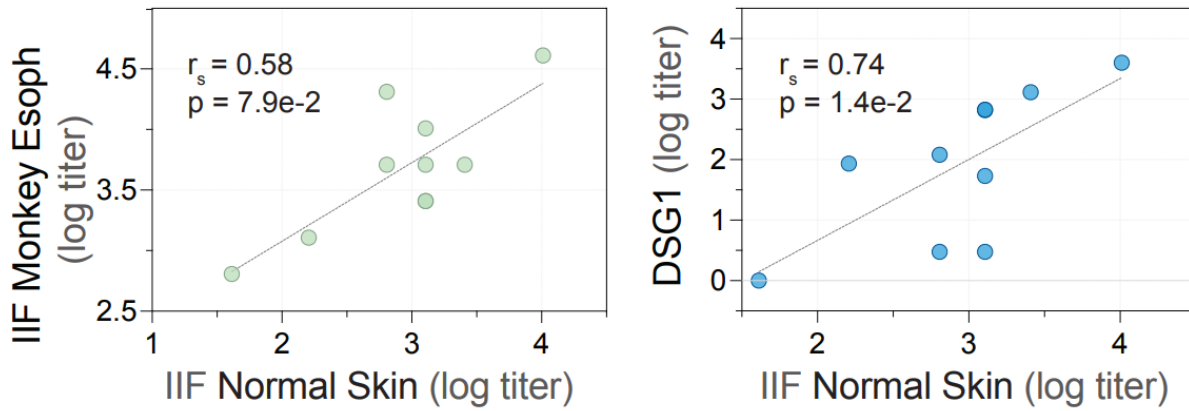


Supplemental Figure 3.2. Age distribution of control group and pemphigus vulgaris patients.

CTRL, control; PV, pemphigus vulgaris.

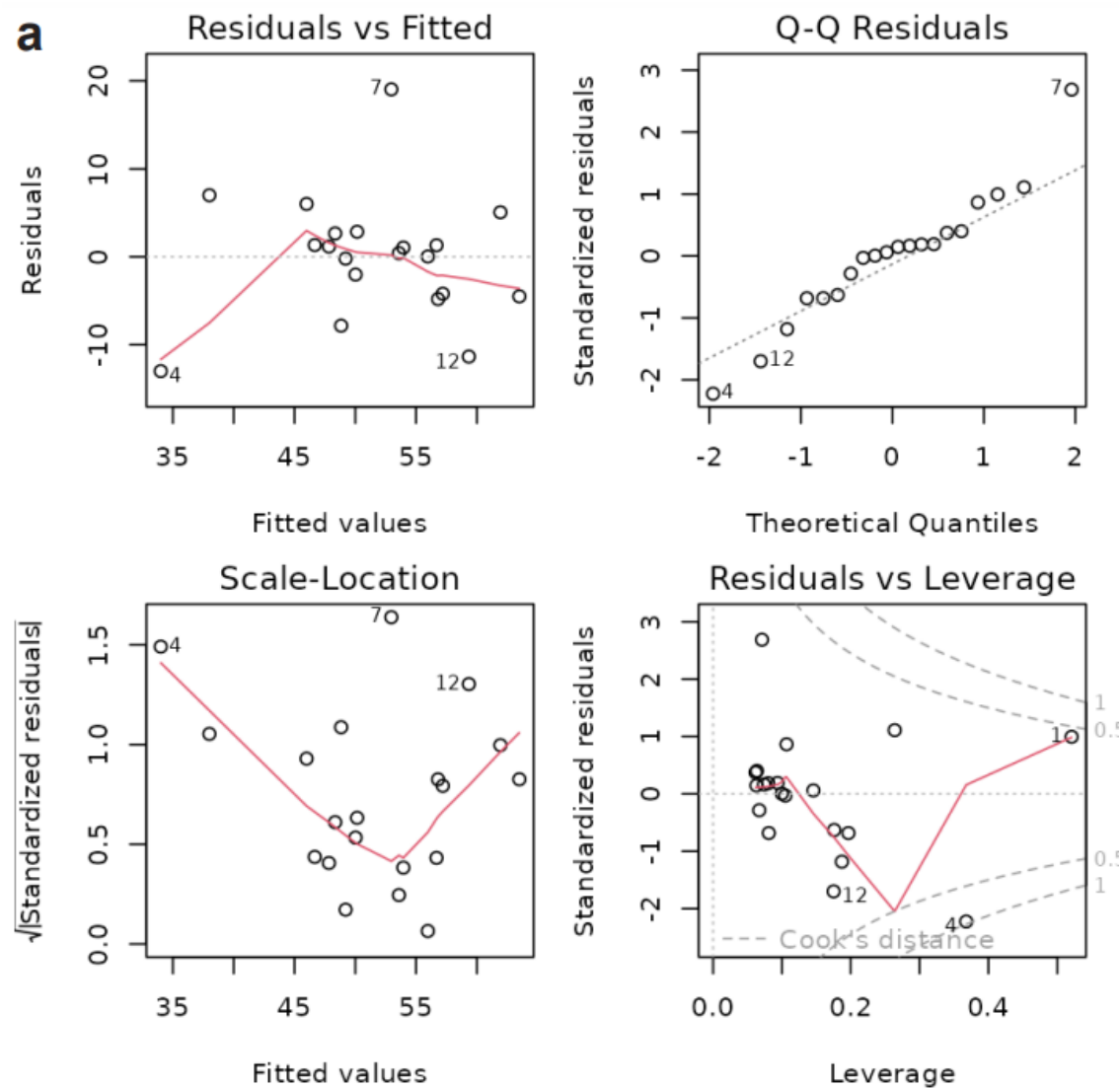


Supplemental Figure 3.3. Similarities and Differences in the Pemphigus and Primary Biliary Cirrhosis (PBC) Serum Glycome. MRM glycoproteomic profiling was performed on serum samples from PBC ($n = 20$) and pemphigus ($n = 10$). Differentially expressed peptides and glycopeptides are shown.



Supplemental Figure 3.4. Surrogate markers of disease severity correlate with each other.

Indirect immunofluorescence (IIF) autoantibody titers to normal skin positively correlate with IIF antibody titers to monkey esophagus and desmoglein-1-specific autoantibody titers.



b

Variance inflation factors (VIF)

lgG1 5410 1.11

lgG2 3410 1.11

Residuals are normally distributed

p-value = 0.28

Residuals have constant variance

p-value = 0.77

Supplemental Figure 3.5. Diagnostic plots for age prediction linear regression model. (a)

Residuals vs. fitted model, assessing linearity; residuals vs. leverage model, assessing overly

influential points; scale-location model, assessing homoscedastic data. **(b)** Variance inflation factors assessing severity of the multicollinearity of the model.

SUPPLEMENTAL TABLES

Supplemental Table 3.1. Ig glycovariants associated with the pemphigus vulgaris. Sera from pemphigus vulgaris patients was compared to healthy controls. Multiple linear regression, p-values and false discovery rates (FDRs) are shown.

Immunoglobulin	Disease p-value	Age p-value	Gender p-value	Disease FDR	Age FDR	Gender FDR
IgG1 g:3410	8.4e-13	5.9e-01	7.9e-02	7.7e-11	1.0	1.0
IgG2 g:3410	7.5e-12	2.8e-01	3.9e-01	6.7e-10	1.0	1.0
IgG3/4 g:3410	1.3e-07	6.3e-01	7.4e-01	1.0e-05	1.0	1.0
IgG1 g:4410	9.9e-12	8.6e-01	1.9e-02	8.7e-10	1.0	1.0
IgG1 g:4411	7.4e-11	3.4e-01	1.7e-03	6.4e-09	1.0	1.3e-01
IgG1 g:4500	4.6e-12	7.6e-01	2.6e-02	4.2e-10	1.0	1.0
IgA2_205 g:5411	7.2e-12	1.5e-01	5.4e-01	6.5e-10	1.0	1.0
IgA1/IgA2 ratio	9.2e-11	5.9e-01	6.0e-02	7.9e-09	1.0	1.0

Supplemental Table 3.2. Pemphigus vulgaris samples set (age, sex, and autoantibody titers).

DSG 1 and DSG 3- DSG-specific autoantibody titers determined by ELISA Monkey esophagus (ME) and normal skin (NS) autoantibody titers determined by indirect immunofluorescence.

Sex	Age	DSG 1	DSG 3	ME	NS
M	67	660	1	2560	1280
F	58	85	275	1280	160
M	56	4000	1	40960	10240
F	21	670	1	2560	1280
F	45	120	800	20480	640
F	41	0	76	640	40
F	72	1300	5	5120	2560
M	54	2	820	10240	1280
M	53	53	1500	5120	1280
F	59	2	310	5120	640

Supplemental Table 3.3. List of monitored transitions in MRM method based on the previously validated methods (Hong et al., 2013, Hong et al., 2015). Glycan IDs correspond to numbers of Hex-HexNAc-Fuc-Neu5Ac residues.

Protein	Glycopeptide/Peptide	Glycan	Precursor Ion (m/z)	Product Ion (m/z)	RT (min)	CE (eV)
IgA2	TPLTAN205ITK	5410	909.8	366.1	6.22	19
IgA2	TPLTAN205ITK	5510	977.5	366.1	6.36	25
IgA2	TPLTAN205ITK	4510	923.5	366.1	6.48	18
IgA2	TPLTAN205ITK	5411	1006.8	366.1	6.9	25
IgA2	TPLTAN205ITK	5511	1074.5	366.1	6.94	25
IgA2	TPLTAN205ITK	5412	1103.8	366.1	7.52	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5512	878.9	366.1	7.66	17
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5501	1016.9	366.1	15.95	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5401	976.3	366.1	15.96	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	3500	1117.1	366.1	16	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	4500	1157.6	366.1	16	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5400	1147.3	366.1	16	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5500	1198.1	366.1	16	25
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	4401	943.9	366.1	16.01	20
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	4501	1230.4	366.1	16.01	30
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5402	1292.9	366.1	16.19	30
IgA1, 2	LSLHRPALEDLLLGSEAN144/131LTCTLTGLR	5502	1075.1	366.1	16.4	25
IgA1	213TPLTATLSK221	—	466.3	415.8	9.03	15
IgA1	154DASGVFTWTPSSGK168	—	770.9	475.3	12.73	20
IgA2	141DASGATFTWTPSSGK155	—	756.9	475.3	11.91	25
IgA1, 2	276YLTWASR282	—	448.7	620.3	10.78	12
IgA1, 2	264WLQGSSELPR273	—	607.4	914.5	11.37	18
IgM	YKN46NSDISSTR	5501	851.1	366.1	2.18	14
IgM	YKN46NSDISSTR	4401	1012.8	366.1	2.19	25

IgM	YKN46NSDISSTR	5411	1115.5	366.1	2.19	25
IgM	YKN46NSDISSTR	4311	993.8	366.1	2.21	25
IgM	YKN46NSDISSTR	5511	1183.2	366.1	2.27	25
IgM	YKN46NSDISSTR	5601	901.9	366.1	2.27	14
IgM	YKN46NSDISSTR	5412	909.6	366.1	2.3	14
IgM	YKN46NSDISSTR	5502	923.9	366.1	2.3	19
IgM	GLTFQQN209ASSMCVPDQDTAIR	5411	1100.5	366.1	12.37	19
IgM	GLTFQQN209ASSMCVPDQDTAIR	5511	1151.3	366.1	12.37	25
IgM	GLTFQQN209ASSMCVPDQDTAIR	4511	1110.8	366.1	12.41	24
IgM	GLTFQQN209ASSMCVPDQDTAIR	5412	1173.3	366.1	12.49	23
IgM	GLTFQQN209ASSMCVPDQDTAIR	5512	1224.1	366.1	12.5	25
IgM	STGKPTLYN439VSLVMSDTAGTCY	9200	1058.3	1284.7	13.51	27
IgM	STGKPTLYN439VSLVMSDTAGTCY	8200	1356.6	1284.7	13.53	27
IgM	STGKPTLYN439VSLVMSDTAGTCY	7200	1302.6	1284.7	13.55	27
IgM	STGKPTLYN439VSLVMSDTAGTCY	5200	1194.5	1284.7	13.58	27
IgM	STGKPTLYN439VSLVMSDTAGTCY	6200	1248.5	1284.7	13.58	24
IgM	STGKPTLYN439VSLVMSDTAGTCY	—	1183.1	342.1	14.28	38
IgM	65YAATSQVLLPSK76	—	639.4	331.2	11.28	15
IgM	301FTCTVTHDLPSPK315	—	573	734.9	12.07	15
J Chain	EN71ISDPTSPLR	5401	1048.1	366.1	9.07	25
J Chain	IIVPLNNREN71ISDPTSPLR	5411	1052.8	366.1	12.63	24
J Chain	IIVPLNNREN71ISDPTSPLR	5412	1125.6	366.1	12.8	18
J Chain	47SSEDPNEDIVER58	—	695.3	971.5	5.78	24
J Chain	62IIVPLNNR69	—	469.8	613.3	11.11	12
IgG1	EEQYN180STYR	5410	986.8	366.1	2.86	15
IgG1	EEQYN180STYR	5400	938.1	366.1	2.88	14
IgG1	EEQYN180STYR	4400	884.1	204.1	2.89	14
IgG1	EEQYN180STYR	4410	932.8	204.1	2.89	15
IgG1	EEQYN180STYR	4500	951.7	204.1	2.9	15
IgG1	EEQYN180STYR	3410	878.8	204.1	2.93	13
IgG1	EEQYN180STYR	5510	1054.5	366.1	2.99	17

IgG1	EEQYN180STYR	4510	1000.5	204.1	3.02	16
IgG1	EEQYN180STYR	3510	946.5	204.1	3.04	15
IgG1	EEQYN180STYR	5411	1083.8	366.1	3.1	17
IgG1	EEQYN180STYR	4411	1029.8	204.1	3.17	16
IgG2	EEQFN176STFR	5410	976.1	366.1	6.46	15
IgG2	EEQFN176STFR	4500	941.1	204.1	6.57	15
IgG2	EEQFN176STFR	5510	1043.8	366.1	6.57	17
IgG2	EEQFN176STFR	4410	922.1	204.1	6.59	25
IgG2	EEQFN176STFR	4400	873.4	204.1	6.62	13
IgG2	EEQFN176STFR	4510	989.9	204.1	6.66	15
IgG2	EEQFN176STFR	3410	868.1	204.1	6.73	13
IgG2	EEQFN176STFR	3510	935.8	204.1	6.81	14
IgG2	EEQFN176STFR	5411	1073.1	366.1	7.19	25
IgG2	EEQFN176STFR	4411	1019.1	204.1	7.41	25
IgG3,4	EEQYN227STFR/EEQFN177STYR	4410	927.4	204.1	4.29	14
IgG3,4	EEQYN227STFR/EEQFN177STYR	4510	995.1	204.1	4.42	15
IgG3,4	EEQYN227STFR/EEQFN177STYR	3510	941.1	204.1	4.47	15
IgG3,4	EEQYN227STFR/EEQFN177STYR	5411	1078.4	366.1	4.6	17
IgG3,4	EEQYN227STFR/EEQFN177STYR	4411	1024.5	204.1	4.69	25
IgG1	158FNWYVDGVEVHNAK171	—	839.41	968.48	12.7	30
IgG2	102CCVECPCPAPPVAGPSVFLFPPKPK127	—	970.14	1100.62	14.65	28
IgG2	272TTPPMLDSDGSFFLYSK288	—	636.31	852.4	14.86	28
IgG3	207WYVDGVEVHNAK218	—	472.91	534.278	11.12	9
IgG4	273TTPPVLDSDGSFFLYSR289	—	634.65	425.22	14.82	9
IgG1&2&3&4	DTLMISR	—	418.22	506.28	8.21	9

Chapter IV

Absolute Quantification of Over 600 Bovine Milk Proteins Using a Novel nanoLC-MS/MS Based Bottom-Up Proteomics Approach

Authors: Armin Oloumi¹, Sheryl Joyce B. Grijaldo-Alvarez¹, Cheng-Yu Weng¹, Ryan Schindler¹, Michael Russelle S. Alvarez¹, Xavier Holmes¹, Jonah Hostettler¹, Mariana Barboza¹, Carlito B. Lebrilla¹

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

ABSTRACT

Proteins are a known nutritional factor in human health affecting major cellular processes, transport of nutrients, and immunity. It is essential to determine the proteome composition of common food products. Furthermore, commonly used methods, such as kjeldahl, only measure total protein contents crudely. While recent advances in food proteomics have enabled comprehensive protein identification, absolute quantification of the vast number of proteins remains challenging due to methodological limitations and the absence of commercial standards. Here, we employed untargeted bottom-up proteomic analysis with LC-MS/MS for the absolute quantification of proteins. This approach requires only minimal protein standards, but it can be used to determine absolute protein concentration values for every detected protein with good accuracy. We demonstrated the use of this method on 60 bovine milk samples collected from different brands, regions, and timepoints. Additionally, a small number of protein standards were used to validate the accuracy of calculated values. Overall, the present method produces absolute quantitation of proteins identified in a proteomic analysis.

INTRODUCTION

The quality of proteins in diet is important because they play key roles in growth and nutrition. Proteins are involved in all cellular processes, including transporting nutrients and fighting diseases (1,2). Traditional protein quantification methods such as bicinchoninic acid (BCA) and Bradford assays yield total protein concentration, but fail to measure abundances of individual proteins (3,4). Specifically, for dairy, the common method used measures total nitrogen, such as the Kjeldahl method, (1,2,5) which produces crude values that have been used to defraud consumers (6).

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) instruments have improved significantly in recent years. Development of nanoscale liquid chromatography (nanoLC) and electrospray ionization sources (nano-ESI) are among major contributors to these advances. Additionally, new mass analyzers, such as orbitraps, are more sensitive and have better resolution than previous generations such as time of flight (TOF). These advancements have resulted in improved protein identifications and sequence coverages in bottom-up proteomics (7,8). New techniques for relative quantification of proteins using LC-MS/MS have also expanded. These methods use isotope or isobaric tags to mark the proteins/peptides in different samples for comparison purposes. Tandem mass tag (TMT) and isobaric tags for relative and absolute quantitation (iTRAQ) are examples of these labels (9,10).

Different dairy products contain a diverse range of proteins, with bovine milk alone containing over 900 proteins (11). Absolute quantification of all milk proteins has significant implications in understanding the functional properties of milk (6,7,12). However, current

methods for absolute quantification of proteins are resource demanding and have major limitations. Furthermore, they require high-purity protein standards and costly isotope labeled homologs. Commercial protein standards are not available for every single milk protein, thereby severely limiting the number of proteins that can be quantified (6,8–10).

Here, we propose a new approach for absolute quantification of proteins using nanoscale LC-MS/MS-based bottom-up proteomics. This approach uses a total protein concentration assay, such as BCA, as opposed to using protein standards and calibration curves for every protein. This method takes advantage of the high resolution and sensitivity of nanoscale LC-MS/MS for comprehensive proteome and sequence coverages of proteins. For quantitation, peptide peak areas of a protein are integrated and normalized to peptide peak areas of all proteins, yielding the relative abundances of the protein. The relative abundances inherently take the protein size into account, because larger proteins yield more peptides with protease digestion. Multiplying the relative abundances with the total protein concentrations yield the absolute concentration of that protein in the sample. Molarity measurements of proteins are eliminated with this approach. This method is advantageous, because there is no “true” molecular weight number for a given protein due to the tremendous heterogeneity in proteoforms. The heterogeneity in proteoforms is due to variations any of dozens of types of post-translational modifications (PTMs) on the proteins.

METHODS AND MATERIALS

Bovine Milk Total Protein Concentration Measurements with BCA

BCA analysis was performed on all Bovine milk samples (n=60, collected from different regions) in triplicates using Pierce BCA Protein Assay Kit (Thermo Scientific, cat. 23225) per instructions from the manufacturer. A 10- μ L aliquot of each milk sample was diluted to 400 μ L with water prior to BCA assay. This dilution lowers the protein concentration to the working range of BCA assay and further reduces the potential matrix effects.

Bovine Milk Protein Digestion Protocol

The protein digestion protocol was adapted from a previously described publication (13). Briefly, a 7- μ L aliquot of each milk sample (n=60) was diluted in 193 μ L of 50 mM ammonium bicarbonate in duplicates. The diluted sample was then reduced with 2 μ L of 550 mM dithiothreitol (DTT) in a water bath at 60 °C for 60 min and alkylated with 4 μ L of 450 mM iodoacetamide (IAA) at room temperature for 30 min in the dark. Proteins were digested by adding 2 μ g of trypsin (sequencing grade, Promega) and incubation for 18 hours in a 37 °C water bath. The samples were then placed in a 4 °C fridge for 1 hour after the digestion.

Bovine Milk Protein Standards Digestion Protocol

A 10 μ g/ μ L solution of α -casein (Sigma cat. C6780), β -casein (Sigma cat. C6905), κ -casein (Sigma cat. C0406), α -lactalbumin (Supelco cat. L-045), β -lactoglobulin (Supelco cat. L3908) and lactotransferrin (Supelco cat. L-047) protein standards were prepared in 50 mM ammonium bicarbonate. α -Casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin and lactotransferrin

were mixed at 0.45, 0.19, 0.15, 0.15, 0.05 and 0.01 mass fractions with the final protein amount of 420 µg. The protein standards were digested, cleaned, and dried using the same protocol as the bovine milk samples. The final sample was reconstituted and diluted to different concentrations to obtain calibration curve points for absolute quantitation of α -casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin and lactotransferrin.

Tryptic Digest Clean-up Protocol

The protein digest clean-up protocol was adapted from a previously described publication (14) with a few minor changes in solution volumes and centrifugation speeds. The digests were purified on 25 mg C18 96-well cartridge plates (ThermoFisher Scientific) by centrifugation (Eppendorf 5810) at 1300 RPM in room temperature. The C18 plate was washed successively with three volumes (200 µL for each volume) of 100% acetonitrile (ACN) in 0.1% TFA and equilibrated with four volumes of pure water in 0.1 % TFA., The tryptic digests were loaded on the plate wells, and then washed with three volumes of pure water in 0.1 % TFA, prior to eluting with two volumes of 40 % ACN in 0.1 % TFA and two volumes of 80 % ACN in 0.1 % TFA. The eluted solutions were dried and stored at -20 °C. The samples were reconstituted in 560 µL of pure water prior to LC-MS/MS analysis.

LC-MS/MS Analysis

The LC-MS/MS method was adapted from a previous publication (15) with minor changes in LC and MS parameters. The cleaned tryptic peptide samples were analyzed on an UltiMate™ WPS-3000RS nanoLC system coupled to an Orbitrap Fusion Lumos MS system (ThermoFisher Scientific). A 1 µL aliquot of each sample was injected, and the peptides were separated on an

Acclaim™ PepMap™ 100 C18 LC Column (75 μ m x 150 mm, particle size 2 μ m; ThermoFisher Scientific) at a flow rate of 300 nL/min. Water in 0.1% formic acid, and 80% acetonitrile in 0.1% formic acid were used as solvents A and B, respectively. The gradient for separation started with 4% B and followed by 0-3 min (4- 9% B), 3-6 min (9-12% B), 6-80 min (12-24% B), 80-95 min (24-50 % B), 95-99 min (50-99% B), 99-103 min (99% B) and 103-103.5 min (99-2% B). MS spectra were collected with a mass range of 375–1500 m/z, resolution of 120000, and scan rate of 1.5 sec/spectrum in positive ionization mode. The filtered precursor ions in each MS spectrum were subjected to fragmentation via 30% higher-energy C-trap dissociation (HCD) using nitrogen gas.

Proteomics Identification and Quantitation

The data analysis was adapted from a previous publication (16). Mass spectrometry data were analyzed using the Byos workflow (Protein Metrics). Proteins were identified on the Byonic software (Protein Metrics) based on the Uniprot Bovine proteome database (Swiss-Prot and TrEMBL combined) with a precursor mass tolerance of 20 ppm and fragment mass tolerance of 20 ppm. The parameters used data analysis included C-terminal cleavage by trypsin (K and R cleavage sites) with up to three missed cleavages, and the following peptide modifications: carbamidomethyl at cysteine, oxidation at methionine, deamidation at asparagine and glutamine, acetylation at protein N- terminal, glutamine to pyro-glutamate at N- terminal, glutamate to pyro-glutamate at N-terminal. Protein IDs were filtered at 1% FDR. Each protein was quantified using the Byologic software (Protein Metrics). Relative abundances of proteins were calculated based on their peptides peak areas. For each protein, the peptides peak areas of that protein were divided by peptides peak areas of all identified proteins. Additional protein information

(Molecular Weight, GO (ID), etc.) were extracted from the Uniprot database. Relative abundances of each protein from every sample were multiplied by the total protein concentration value obtained from the BCA assays.

GO Enrichment Analysis

Gene Ontology (GO) enrichment of the identified bovine milk proteins was performed using G:Profiler (<https://biit.cs.ut.ee/gprofiler>) for the *Bos taurus* (cow). Three categories including cellular component (CC), molecular function (MF) and biological process (BP) were used to classify the bovine milk proteins. While focusing on driver terms in the GO analysis, the hypergeometric p-values were recalculated using 'greedy algorithm' search strategy resulting in at least one function being highlighted from every connected component, which ensured that leading terms are presented rather than just the highest significance level. The top 30 pathways focusing on the leading terms related to bovine milk proteins were determined and visualized using Graphpad Prism. The protein bioactivities were classified to be under 5 categories: metabolism, coagulation, system process, response to stimulus and transport using the ancestor chart for GO terms in EMBL-EBI (<https://www.ebi.ac.uk/QuickGO/term>).

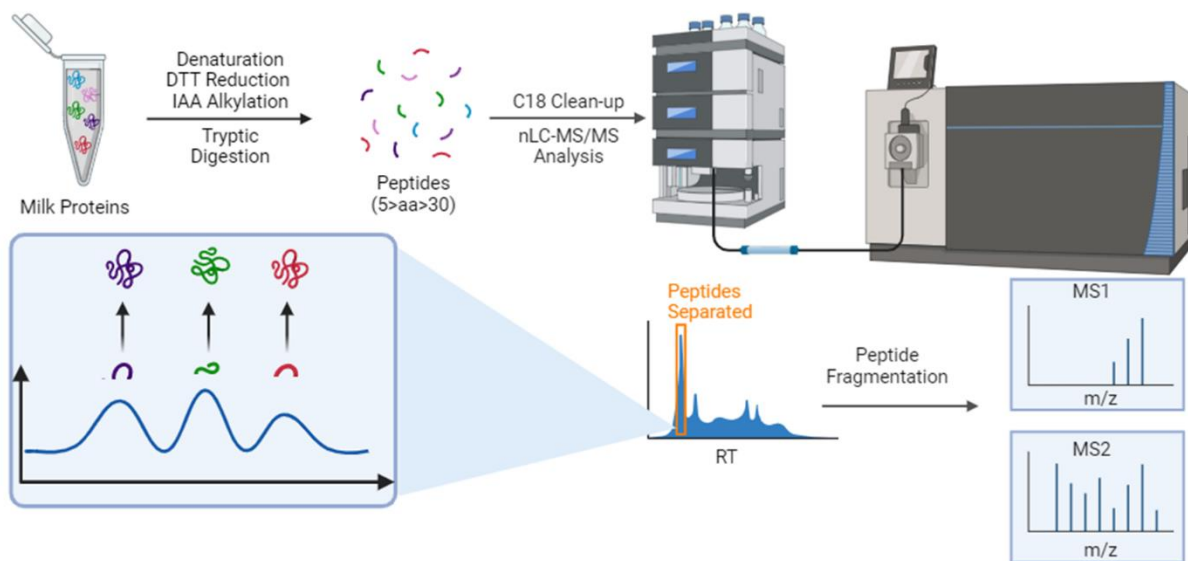
RESULTS AND DISCUSSION

Conventional vs Novel Bovine Milk Protein Quantification

In this study, we developed an untargeted bottom-up LC-MS/MS proteomics method to determine absolute quantification of proteins in bovine milk samples from various commercial

brands, regions and timepoints (**Figure 4.1A**). Every detectable protein was quantified with absolute abundances using BCA assays combined with only a few protein standards (**Figure 4.1B**). This approach eliminated the issues resulting from the lack of availability of protein standards for every detected protein. Protein standards often contain different PTMs compared to the proteins in analytes. These inconsistencies in PTMs often result in inaccurate concentration measurements, because generally only one or two peptides are used for protein quantification, and these peptides may contain PTMs.

(A)



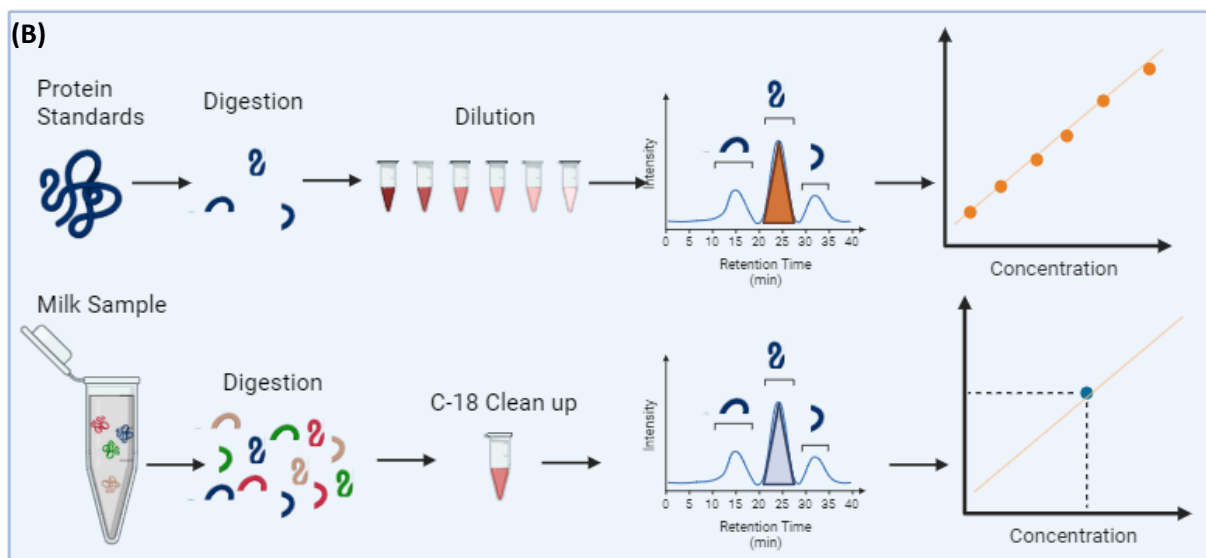


Figure 4.1. (A) Bottom-up proteomics workflow used for protein identification **(B)** Protein quantification using LC/MS with protein standards.

Total protein concentrations were measured in triplicates using the standard BCA assay (**Figure 4.2**). The total protein abundances in conjunction with the relative abundances of individual proteins, as obtained by LC-MS/MS, were then used to calculate absolute concentrations of individual proteins as described in greater detail below. The total protein concentrations measured by BCA ranged from 38 mg/mL to 59 mg/mL, with an average of 49 mg/mL and a coefficient of variation (CV) of 8.8% (**Figure 4.2**). In terms of nutrition, average value is 11.6 g of protein per cup, which is just slightly greater than the commonly reported value of 8 g/cup on the commercial milk labels obtained commonly through the Kjeldahl method.

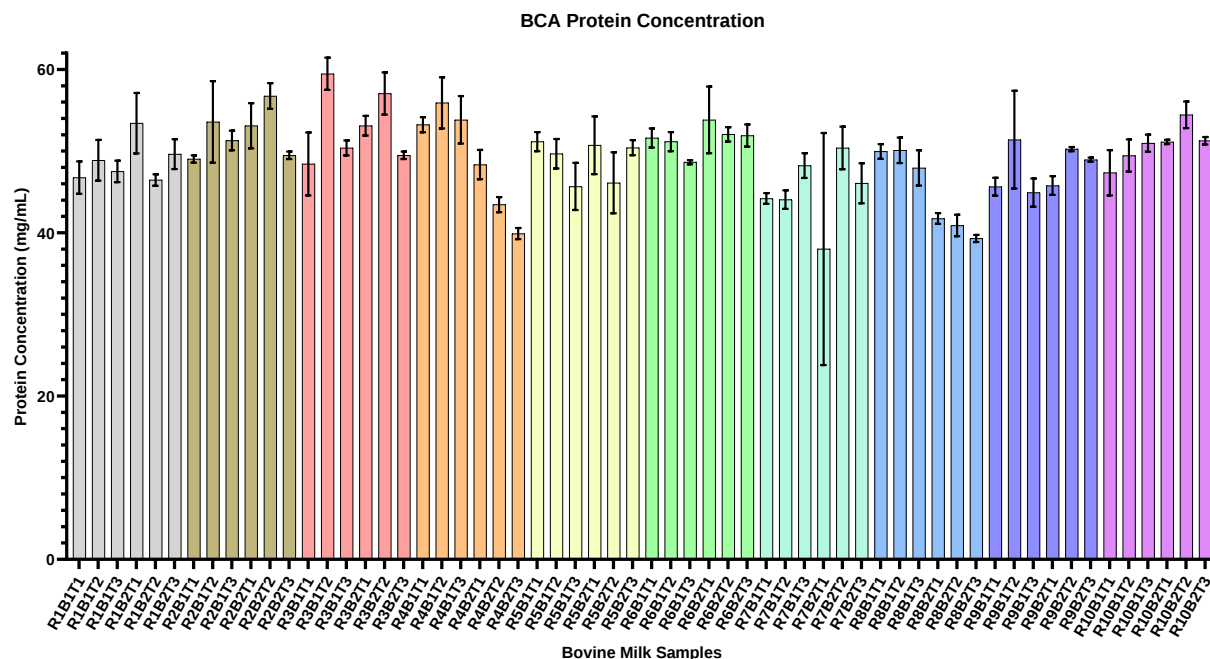


Figure 4.2. BCA Protein Concentration Measurements of Bovine Milk samples (n=60). Total protein concentrations were measured in triplicates (average: 49 mg/mL; range: 38 - 59 mg/mL; value on the labels: 34 - 38 mg/mL); R=region, B=brand, T=timepoint.

After measuring total protein concentration using BCA, the samples were subjected to the proteomics workflow starting with digestion using trypsin. LC-MS/MS analysis of the samples yielded identification of over 600 bovine proteins. To obtain absolute abundance for each protein, the relative abundances of individual proteins were first determined. The relative abundance of a protein was calculated by summing the peak areas of every peptide from that protein and dividing it by the sum of the peak areas of peptides from all identified proteins. The full proteome coverage was obtained by bottom-up proteomics with the sum of the corresponding peptide peak areas for each protein used to determine a relative protein abundance (RPA). By multiplying the RPA of a protein with the total protein concentration (TPC), obtained from BCA, the absolute

concentration of that protein is determined (**Eq. 4.1**). Using this equation the concentration of each protein was determined in grams per volume rather than moles per volume. The size of the protein scales with the number of peptides. Thus, large proteins represented by more peptides gives larger peak areas compared to smaller proteins with less peptides.

$$\text{Relative Protein Abundance (RPA)} = \frac{\text{Sum of Peak Areas of Peptides from **One Protein**}}{\text{Sum of Peak Areas of Peptides from **All Identified Proteins**}}$$

$$\text{Absolute Concentration } \left(\frac{g}{\text{volume}} \right) = \text{Relative Protein Abundance (RPA)} \times \text{Total Protein Concentration (TPC)}$$

Eq. 4.1. Equations for relative protein abundance (RPA) and absolute concentration calculations of proteins.

Proteomic analyses of the bovine milk samples in duplicates yielded over 600 bovine proteins, which were quantified using this method (summarized in **Table 4.1**). Absolute concentrations for these proteins are reported in µg/cup to provide the amount of each protein in one serving (one cup or 236.5 mL). Concentrations of individual proteins ranged from 0.002 to 4,000,000 µg/cup. Of over 600 proteins, the absolute concentrations of only 19 were previously reported (17). The most abundant six proteins corresponded to 94.3% of the total protein abundances. The most abundant and the next most abundant (6+49 proteins) represented 99.7% of the total abundance. A large fraction of the quantified proteins (327+252 proteins) makes up less than 0.29% of the total abundance.

Table 4.1. Absolute quantification of 634 proteins from 60 bovine milk samples with concentrations ranging from 0.002 to 4,000,000 µg/cup.

Number of Proteins	Concentration Range (µg/cup)
6	500,000 - 4,000,000
49	1,000 – 120,000
327	5 – 1,000
252	0.002 - 5

Validation of Absolute Protein Concentrations

To validate the accuracy of the absolute abundances obtained by above method, a milk protein standard mixture of α -casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin and lactotransferrin was prepared with 45, 19, 15, 15, 5 and 1 mg/mL concentrations, respectively. This mixture was analyzed using the same proteomic method as the milk samples. Over 150 proteins were identified in this mixture of six protein standards, highlighting a major issue with gravimetric measurements of protein standards for absolute quantification. Protein standards may contain protein impurities that cause inaccuracies in the mass measurements and thus protein concentrations. For relative protein abundance (RPA) calculations, the peptide peak areas of each protein standard were divided by the total peptide peak areas of all six protein standards. The relative protein abundances obtained through this analysis yielded absolute concentrations of 45, 8.4, 15.6, 22.1, 6.4 and 2.4 mg/mL for α -casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin, and lactotransferrin, respectively. The obtained concentrations for α -casein, κ -casein, α -lactalbumin are in close agreements with the actual values based on gravimetric measurements with minor differences of 0%, 4% and 29% respectively (**Figure 4.3**). Furthermore, the differences observed in protein concentrations of β -casein, β -lactoglobulin, and lactotransferrin were 56%, 48% and 138%, respectively (**Figure 4.3**). These differences may be due to the presence of protein impurities in each of the standards and inaccuracies in gravimetric measurements as the result. Additionally, the number of lysine and arginine amino acids of each protein may have also contributed to these differences, as those are the cleavage sites for trypsin protease. Furthermore, greater numbers of lysine and arginine residues in peptides may result

in greater ionization efficiencies and better detection of peptides as they get positively charged through electrospray ionization.

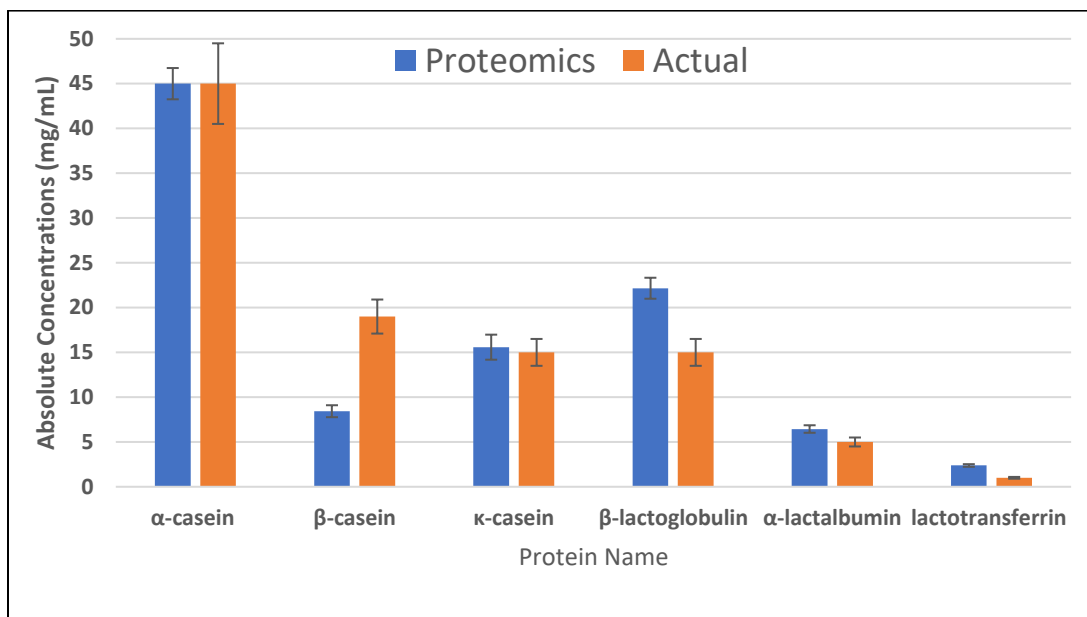


Figure 4.3. Bar graph representing the concentrations of proteins standards obtained by RPA and TPC (proteomics) vs. actual (gravimetric method).

The reproducibility of relative proteomic abundance measurements was assessed by injecting different amounts of peptides digests into the instrument and comparing the proteomic abundances of protein standards with different injection amounts of peptides. The digested protein standards mixture went through dilutions and different amounts of peptides ranging from 0.08 μg to 0.67 μg were injected and analyzed by LC-MS/MS. With different injection amounts, the coefficient of variations in relative protein abundances (RPA) of α -S1-casein, α -S2-casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin and lactotransferrin were 3.7%, 5.6%, 14.6%, 4.9%, 1.9%, 7.4% and 9.4%, respectively (**Figure 4.4**). These CV values are relatively small considering different amounts of peptides were injected in each run. Often injections of less peptides may skew the proteomic abundances due

to lack of detection of peptides from lower abundance proteins, but that is not the case with this method.

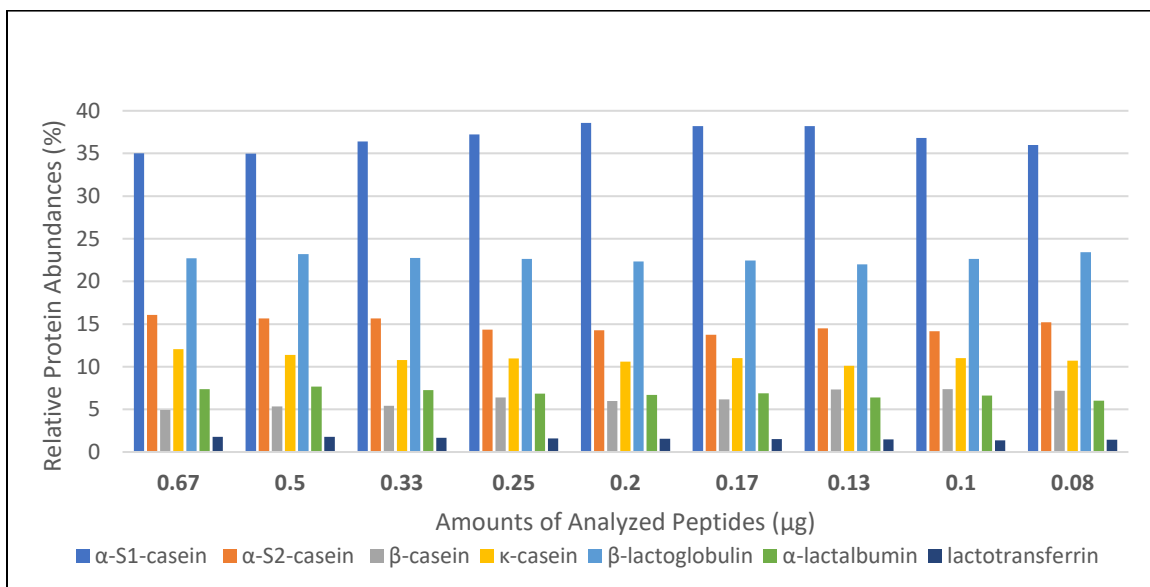


Figure 4.4. Bar graph representing the measured RPA of protein standards by LC-MS/MS with injection of different peptide amounts (0.08 – 0.67 μg).

The linearity of peptides peak areas for each protein standard was also demonstrated with different injection amounts, and external calibration curves were made for the protein standards (**Supplementary Figure 4.1**). These calibration curves were used to determine concentrations of α -casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin, and lactotransferrin in the actual milk samples. Calculated concentrations using the calibration curves were compared to the ones obtained by RPA combined with TPC method, as well as the concentrations reported in the literature (17) (**Table 4.2**). RPA combined with TPC method yielded absolute concentrations of 13.8, 3.8, 2.0 and 0.27 (g/L) for α -S1-casein, κ -casein, α -lactalbumin and lactotransferrin, respectively. These concentrations fell within the range of concentrations reported in the literature for α -S1-casein (10-15 g/L), κ -casein (2.5-

4 g/L), α -lactalbumin (0.74-2.06 g/L) and lactotransferrin (0.09-0.33 g/L) (17) (**Table 4.2**). Additionally, RPA combined with TPC method yielded absolute concentrations of 6.4, 5.8 and 7.5 for α -S2-casein, β -casein and β -lactoglobulin, respectively, and these values fell outside of the concentration ranges reported in the literature for α -S2-casein (2.6-4.02 g/L), β -casein (9.3-11.33 g/L) and β -lactoglobulin (3.2-4.12 g/L) (17) (**Table 4.2**). It is important to note these slight differences may be due to the differences in sources of milk samples and the standards used in other studies.

External calibration curve method yielded the concentration of 2.6 g/L for α -S2-casein, which fell within the literature concentration range of 2.6–4.02 g/L (17) (**Table 4.2**). However, the calibration curve method also yielded concentrations of 5.4, 8, 2.3, 2.2, 0.62 and 0.062 g/L for α -S1-casein, β -casein, κ -casein, β -lactoglobulin, α -lactalbumin and lactotransferrin, respectively, and these values were slightly outside of the concentration ranges reported in the literature for α -S1-casein (10-15 g/L), β -casein (9.3-11.33 g/L), κ -casein (2.5-4 g/L), β -lactoglobulin (3.2-4.12 g/L), α -lactalbumin (0.74-2.06 g/L) and lactotransferrin (0.09-0.33 g/L) (17) (**Table 4.2**). Although these values appear to not agree with the literature concentrations as well as the concentrations obtained by RPA and TPC, majority of the differences are within 10% of the literature reported values, which is an acceptable confidence interval range for absolute concentration measurements. The differences may be due to the differences in methods, as well as the sample sources and the purity of the standards (17) (**Table 4.2**).

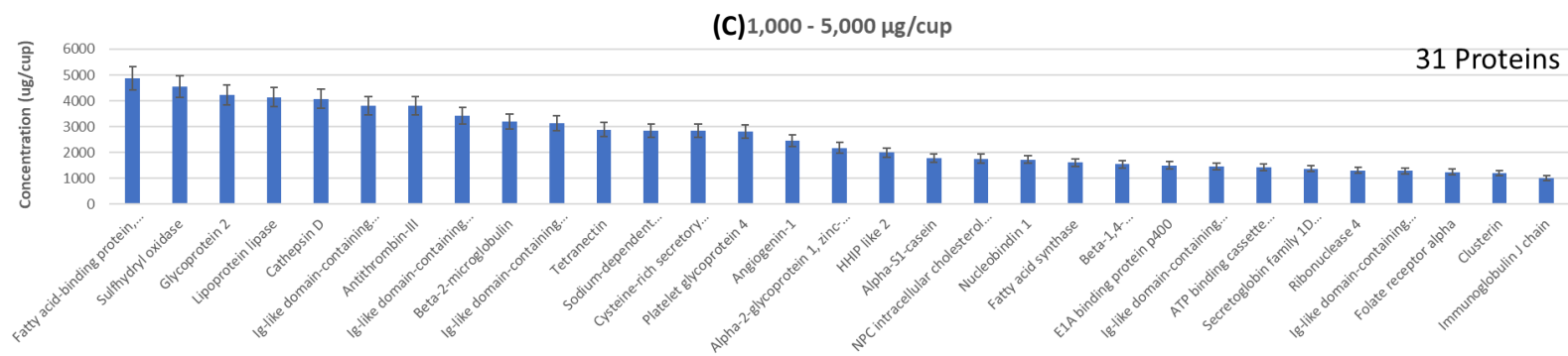
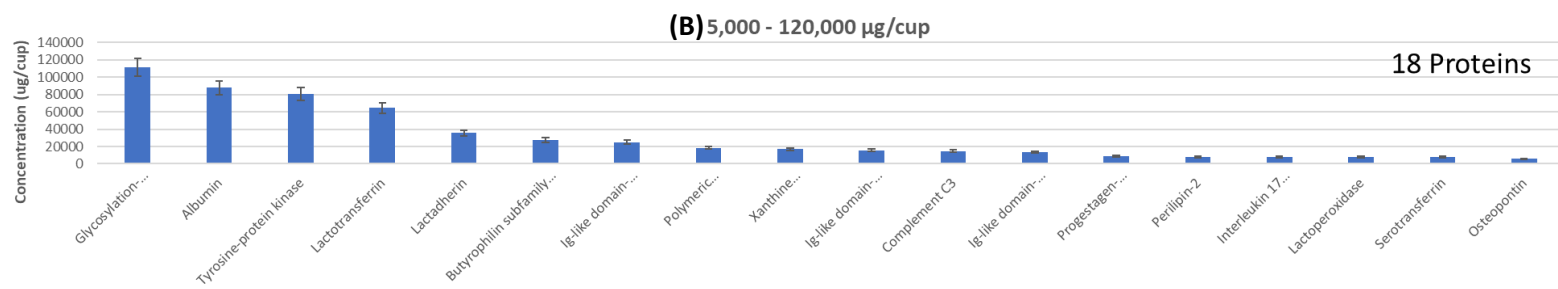
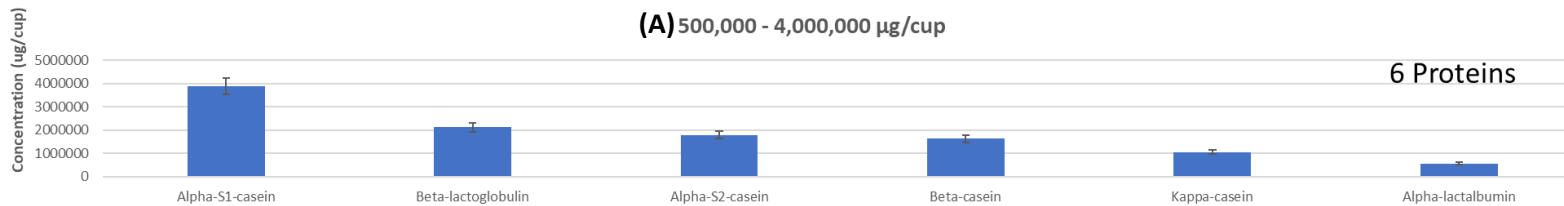
Table 4.2. Comparisons of concentrations of major milk proteins obtained by Calibration Curve Method, Relative Protein Abundance (RPA) combined with Total protein Concentration (TPC) Method, and the concentrations reported in the literature.

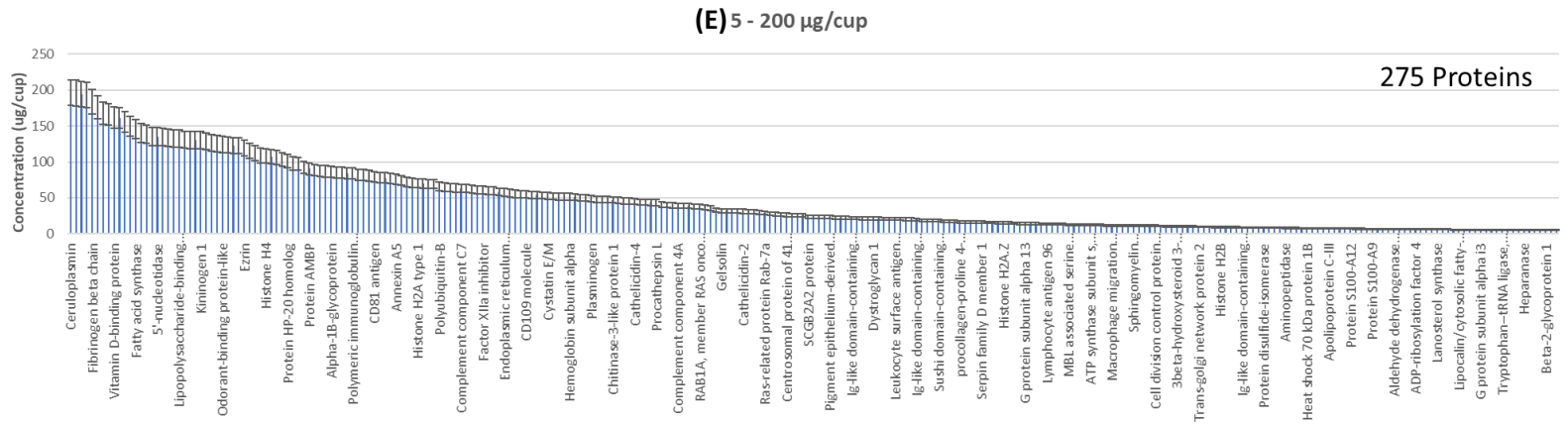
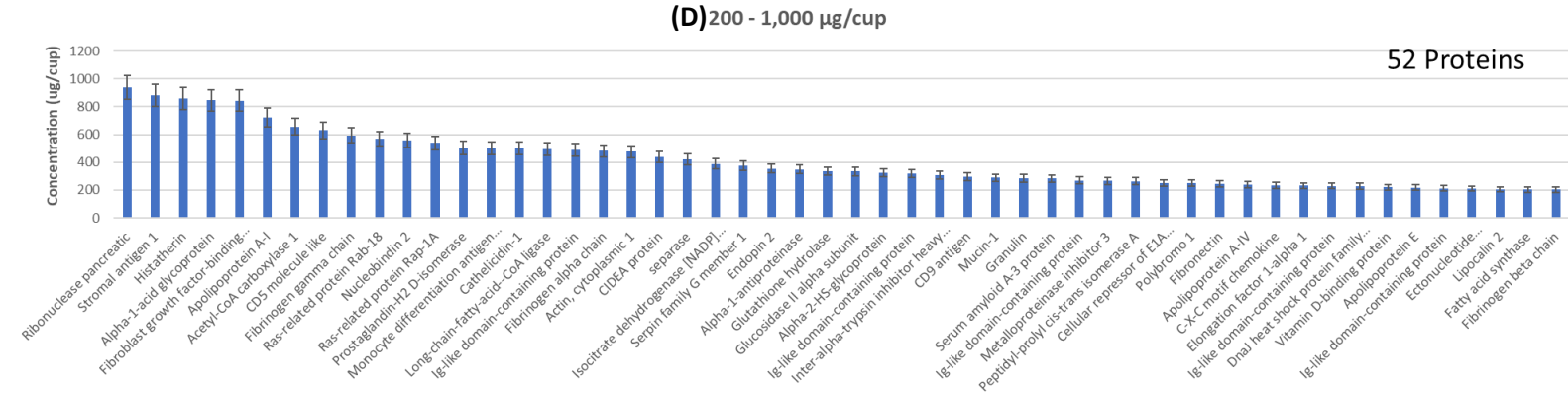
Protein Name	Protein ID	Calibration Curve Method (g/L)	RPA and TPC (g/L)	Literature Values (17) (g/L)
ALPHA-S1-CASEIN	P02662	5.4	13.8	10 - 15
ALPHA-S2-CASEIN	P02663	2.6	6.4	2.6 - 4.02
BETA-CASEIN	P02666	8	5.8	9.3 - 11.33
KAPPA-CASEIN	P02668	2.3	3.8	2.5 - 4
BETA-LACTOGLOBULIN	P02754	2.2	7.5	3.2 – 4.12
ALPHA-LACTALBUMIN	P00711	0.62	2.0	0.74 – 2.06
LACTOTRANSFERRIN	P24627	0.062	0.27	0.09 – 0.33

Total protein concentration (TPC) measurements by BCA were also validated by the standard addition method. BCA assay may yield inaccurate total protein concentrations due to matrix effects, such as presence of lactose and fats in milk. For standard addition, known amounts of the protein standard mixture were spiked in a milk sample (**Figure S2**). Standard addition method yielded a total protein concentration of 37.7 mg/mL for the milk sample pool, while the conventional BCA, which uses an external calibration curve, yielded 48.5 mg/mL. Conventional BCA may be overestimating the total protein concentrations by 29% based on the standard addition results. The protein standard mixture used for standard additions only contained 6 protein standards, and that may have also contributed to the minor inconsistencies in measured concentrations by conventional BCA and standard addition methods.

Absolute Concentrations of Bovine Proteins in Milk

The proteins with their average absolute concentrations are organized from the most abundant to the least in **Figure 4.5**, and concentrations of top 55 most abundant proteins in each milk sample are presented in **Figure 4.6**. Absolute concentration comparisons of these 55 high abundant proteins show the milk proteome composition is conserved across different milk samples (**Figure 4.6**). Each sample is composed mostly of prominent milk proteins (α -S1-Casein, β -Lactoglobulin, α -S2-Casein, β -Casein, κ -Casein, α -Lactalbumin), which make up about 95% of milk proteome in terms of the quantity, and the concentrations of these proteins ranged from 500,000 to 4,000,000 $\mu\text{g}/\text{cup}$ (**Figure 4.5A**). The remaining 628 quantified milk proteins in this study make up about 5% of milk proteome in terms of the quantity and the concentrations of these proteins ranged from 0.002 to 120,000 $\mu\text{g}/\text{cup}$ (**Figure 4.5B-G**). Among this group of proteins, there were 579 proteins with relatively low concentrations of 0.002-1,000 $\mu\text{g}/\text{cup}$ (**Figure 4.5D-G**). These low abundant proteins are typically overlooked in bovine milk protein absolute quantification studies.





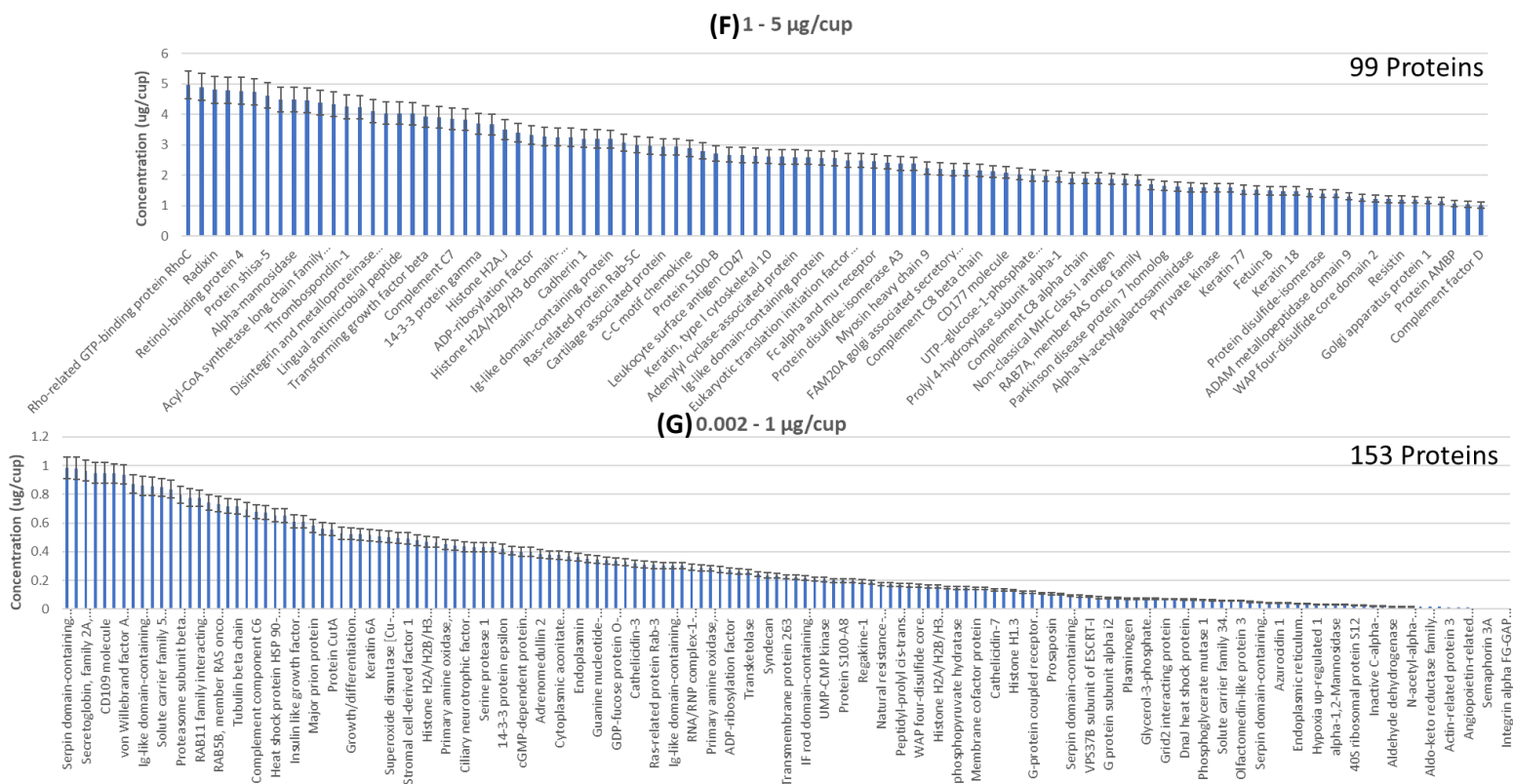


Figure 4.5. Absolute quantification of 634 proteins from 60 bovine milk samples with concentrations ranging from **(A)** 500,000 - 4,000,000 $\mu\text{g}/\text{cup}$ – 6 proteins, **(B)** 5,000 - 120,000 $\mu\text{g}/\text{cup}$ – 18 proteins, **(C)** 1,000 - 5,000 $\mu\text{g}/\text{cup}$ – 31 proteins, **(D)** 200 - 1,000 $\mu\text{g}/\text{cup}$ – 52 proteins, **(E)** 5 - 200 $\mu\text{g}/\text{cup}$ – 275 proteins, **(F)** 1 - 5 $\mu\text{g}/\text{cup}$ – 99 proteins, **(G)** 0.002 - 1 $\mu\text{g}/\text{cup}$ – 153 proteins.

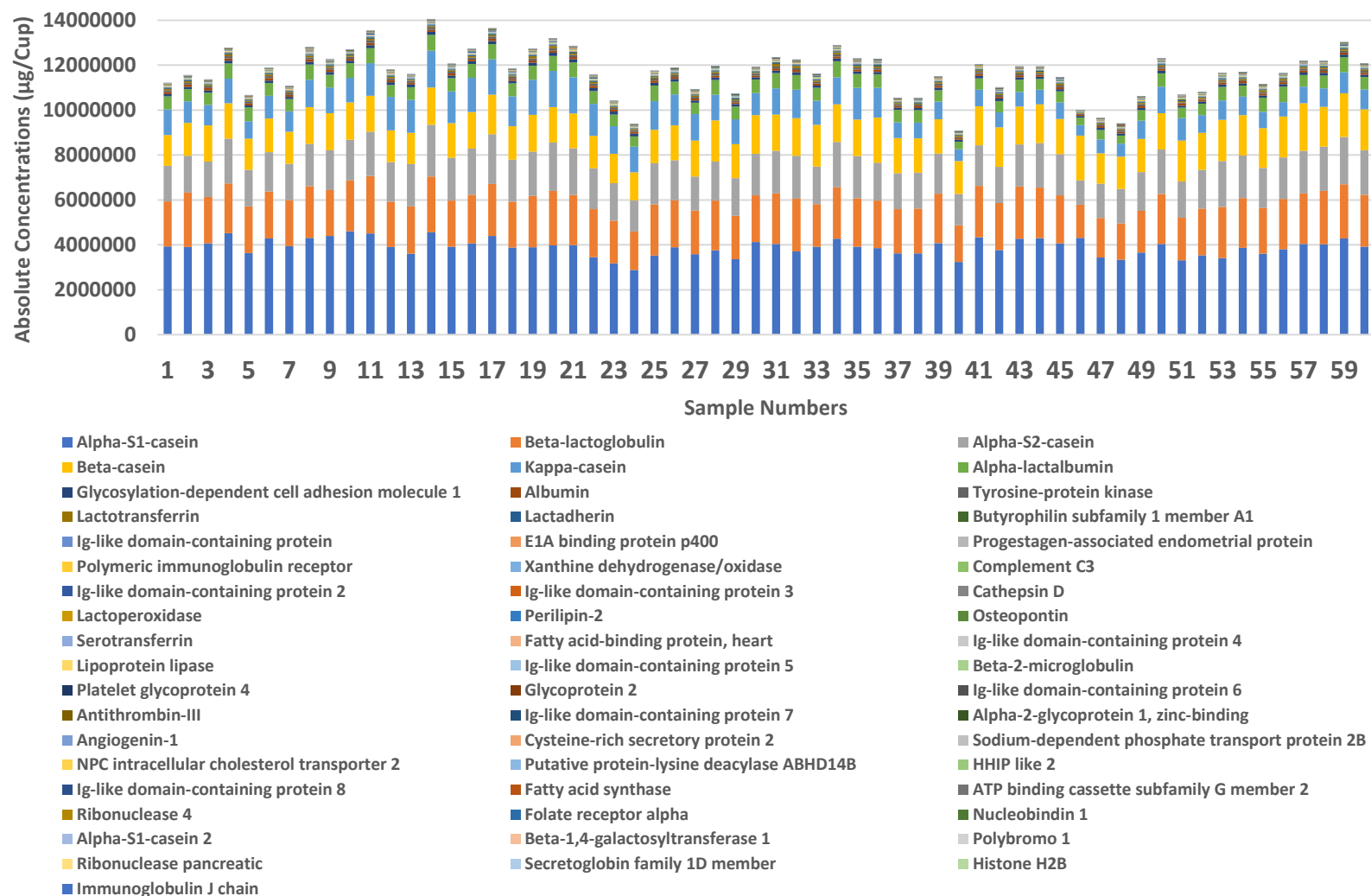


Figure 4.6. Absolute concentrations comparisons of 55 high-abundant proteins in all bovine milk samples (n=60).

Many of the quantified bovine milk proteins have biological activities, which are described in greater details below. Osteopontin (5,600 µg/cup) and lactotransferrin (64,500 µg/cup) are among these biologically active proteins. Osteopontin (P31096) is a phosphoprotein that belongs to the small integrin-binding ligand N-linked glycoprotein (SIBLING) family. Studies have linked osteopontin with bone remodeling, as its expression correlates with cartilage-to-bone transitions during chondrocyte maturation (18,20). Also, osteopontin is a potential biomarker for early diagnosis of postmenopausal osteoporosis (20). Osteopontin concentration in all samples is presented in **Figure 4.7**. The concentration ranged from 1,500 to 8,300 µg/cup in the samples with an average of 5,600 µg/cup. Lactotransferrin, also known as lactoferrin (P24627), is a major iron-binding glycoprotein that belongs to the transferrin (TF) family. Lactotransferrin is abundant in exocrine fluids such as breast milk and mucosal secretions, and it has antimicrobial properties against *E. coli* and *P. aeruginosa* (18,19,21). Lactotransferrin concentrations ranged from 33,000 to 107,000 µg/cup in samples with an average of 64,500 µg/cup (**Figure 4.8**).

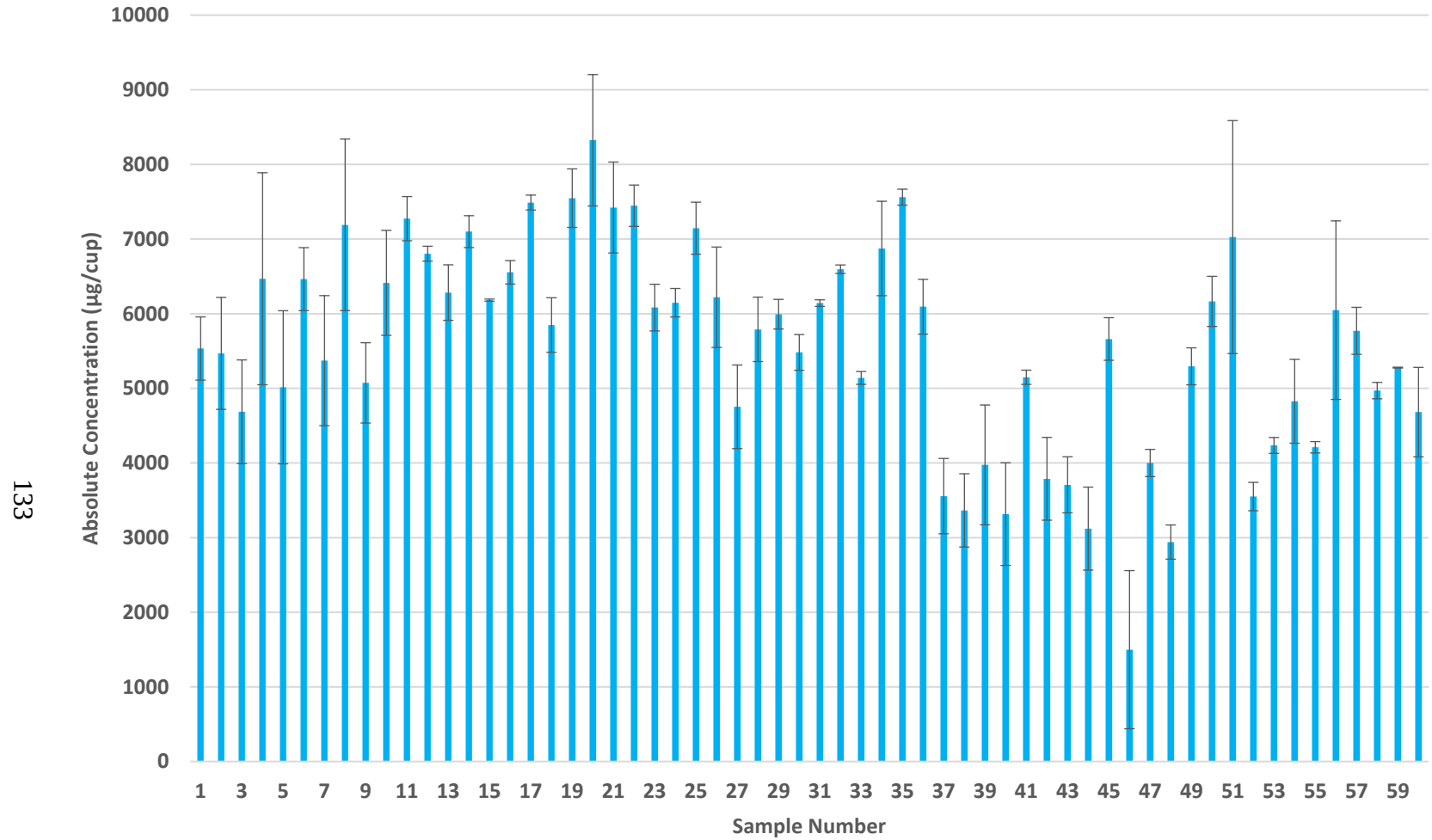


Figure 4.7. Absolute concentrations of Osteopontin in bovine milk samples (n=60) ranged from 1,500 to 8,300 µg/cup.

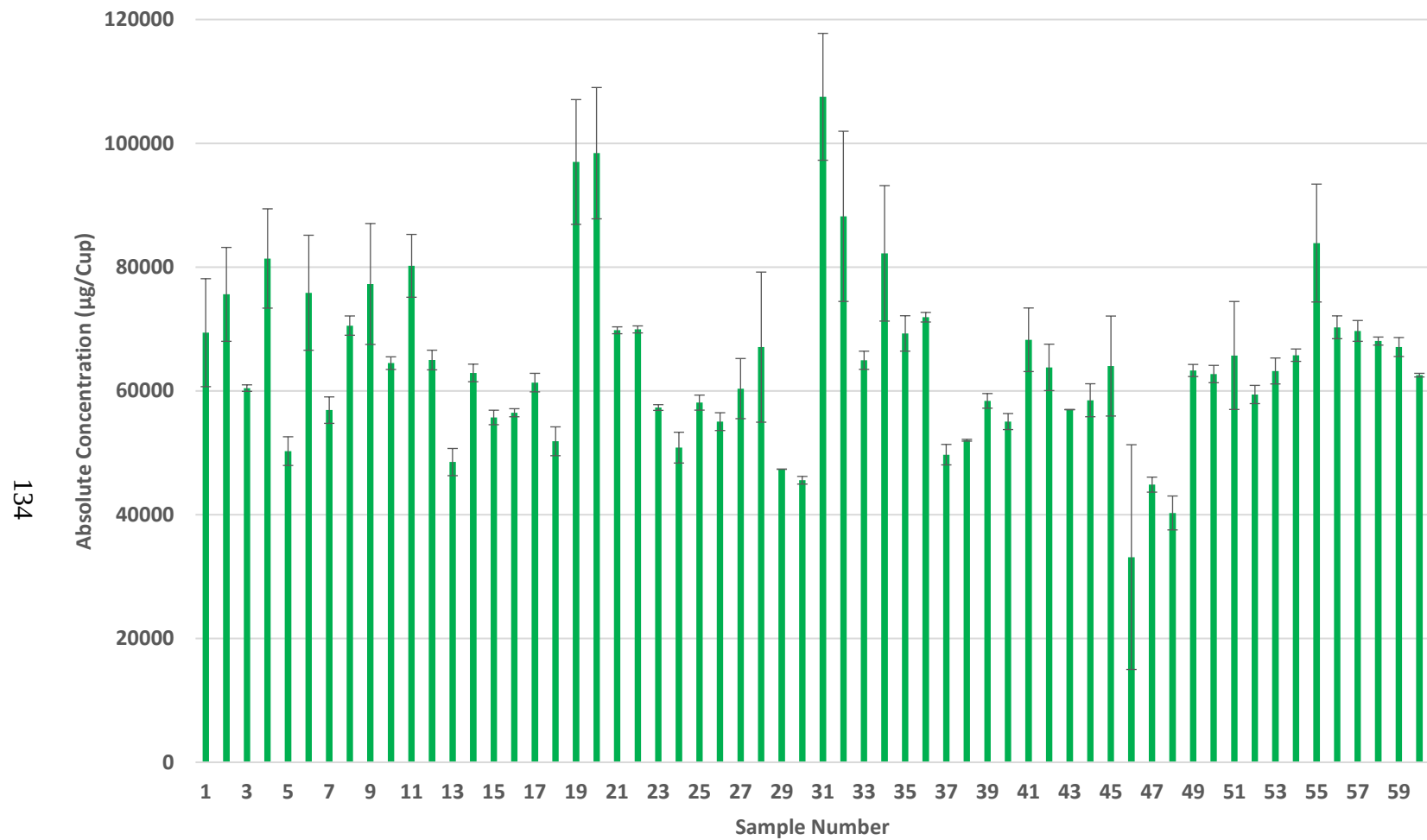


Figure 4.8. Absolute concentrations of Lactotransferrin in bovine milk samples (n=60) ranged from 33,000 to 107,000 µg/cup

Quantified Bovine Milk Proteins Have Potential Health Benefits

Bioactivity and potential health benefits of quantified proteins are summarized in **Table 4.3**. Many of the proteins have antioxidant and antibacterial properties. Biological roles of quantified proteins were further explored through gene ontology analysis (GO). The proteins were classified based on the three GO categories of cellular component (CC), molecular function (MF) and biological process (BP) (**Figure 4.9A**). Most of the proteins belong to the endomembrane system, cytoplasmic vesicle, and endoplasmic reticulum (ER) cellular components, and they affect many molecular functions, such as peptidase regulator activity, GTPase activity, signal receptor binding, antioxidant activity, immunoglobulin receptor binding, and G-protein beta/gamma-subunit complex binding.

Milk proteins are involved in many biological processes, most notably immune system process, immune response, humoral immune response, defense response, transport, immune effector process, inflammatory response, response to wounding and wound healing (**Figure 4.9A**). Based on the GO, bovine milk proteins appear to be heavily involved with immunity. The top 30 pathways focusing on the leading terms related to bovine milk proteins are presented in **Figure 4.9B**. The five categories of metabolism, coagulation, system process, response to stimulus and transport were among the most prominent bioactivities of bovine milk proteins.

Table 4.3. Select quantified bovine proteins with known health benefits.

Protein Name	Protein ID	Concentration Range (µg/cup)	Health Benefits	Reference
ALPHA-S1-CASEIN	P02662	2,639,809	Antioxidant, antimicrobial, apoptosis/anticarcinogenic, bone growth, anti-inflammatory and growth	(19,22–25)
BETA-LACTOGLOBULIN	P02754	1,438,606	Antioxidant, anti-inflammatory, antidiabetic, antimicrobial and anticarcinogenic	(19,22,23,26,27)
ALPHA-S2-CASEIN	P02663	1,233,197	Antioxidant, antimicrobial, anticarcinogenic, antihypertensive, antidiabetic	(19,22,23,28–30)
BETA-CASEIN	P02666	570,573	Bone growth, antioxidant, antihypertensive, and anti-inflammatory	(19,22,23,31–34)
KAPPA-CASEIN	P02668	388,589	Antioxidant, cytotoxic, antimicrobial, anti-carcinogenic, anti-Inflammatory, anticarcinogenic	(19,22,23,35,36)
ALPHA-LACTALBUMIN	P00711	383,037	Antioxidant	(19)
GLYCOSYLATION-DEPENDENT CELL ADHESION MOLECULE 1	P80195	75,733	Antimicrobial	(19)
TYROSINE-PROTEIN KINASE	A0A3Q1MYA3	54,691	Maternal/Prenatal Challenges, Energy for Everyday Living/Work and Immunity	(37)
LACTOTRANSFERRIN	P24627	64,500	Inflammation	(19)
ALBUMIN	P02769, A0A140T897	38,450	Antioxidant	(19)
LACTADHERIN	Q95114	23,990	Immunity	(19)
GTPASE, IMAP FAMILY MEMBER 7		18,183	Cancer Prevention	(38)
POLYMERIC IMMUNOGLOBULIN RECEPTOR	P81265	12,415	Inflammation, antimicrobial and immunity	(19)
BUTYROPHILIN SUBFAMILY 1 MEMBER A1	F1MZQ4, P18892	11,273	Antimicrobial	(39)
XANTHINE DEHYDROGENASE/OXIDASE	P80457, F1MUT3	10,269	Antimicrobial	(19)
COMPLEMENT C3	Q2UVX4	10,047	Anti-sepsis	(19)
PROGESTAGEN-ASSOCIATED ENDOMETRIAL PROTEIN	A0A3Q1LYE8	6,073	Reproduction/Fertility Challenges	(19)
LACTOPEROXIDASE	P80025	5,380	Inflammation and Antimicrobial	(19)
INTERLEUKIN 17 RECEPTOR E LIKE	G3N0X3	4,999	Antimicrobial and Immunity	(40)
OSTEOPONTIN	P31096	5,600	Bone growth and Fraction repair	(19,20)
ANGIOGENIN-1	P10152	1,668	Immunity	(19)
ALPHA-2-GLYCOPROTEIN 1, ZINC-BINDING	A0A452DK44	1,466	Immunity	(19)
RIBONUCLEASE 4	P15467	886	Antimicrobial and immunity	(19)
CLUSTERIN	P17697	853	Inflammation and antimicrobial	(19)
IMMUNOGLOBULIN J CHAIN	Q3SYR8	684	Immunity	(19)
HISTATHERIN	C6KGD8	581	Antimicrobial	(19)
ALPHA-1-ACID GLYCOPROTEIN	Q3SZR3	571	Inflammation	(19)
APOLIPOPROTEIN A-I	P15497	487	Inflammation	(19)
ACETYL-COA CARBOXYLASE 1	Q9TTS3	455	Energy	(19)
NUCLEOBINDIN 2	F6Q8A8	380	Bone growth	(19)
RAS-RELATED PROTEIN RAP-1A	P62833	366	Growth	(19)
LONG-CHAIN-FATTY-ACID--COA LIGASE	A0A3Q1LWT4	366	Energy	(19)
MONOCYTE DIFFERENTIATION ANTIGEN CD14	Q95122	340	Immunity	(19)
CATHELICIDIN-1	P22226	339	Antimicrobial and immunity	(19)
SERPIN FAMILY G MEMBER 1	E1BMJ0	252	Inflammation, immunity and antioxidant	(19)

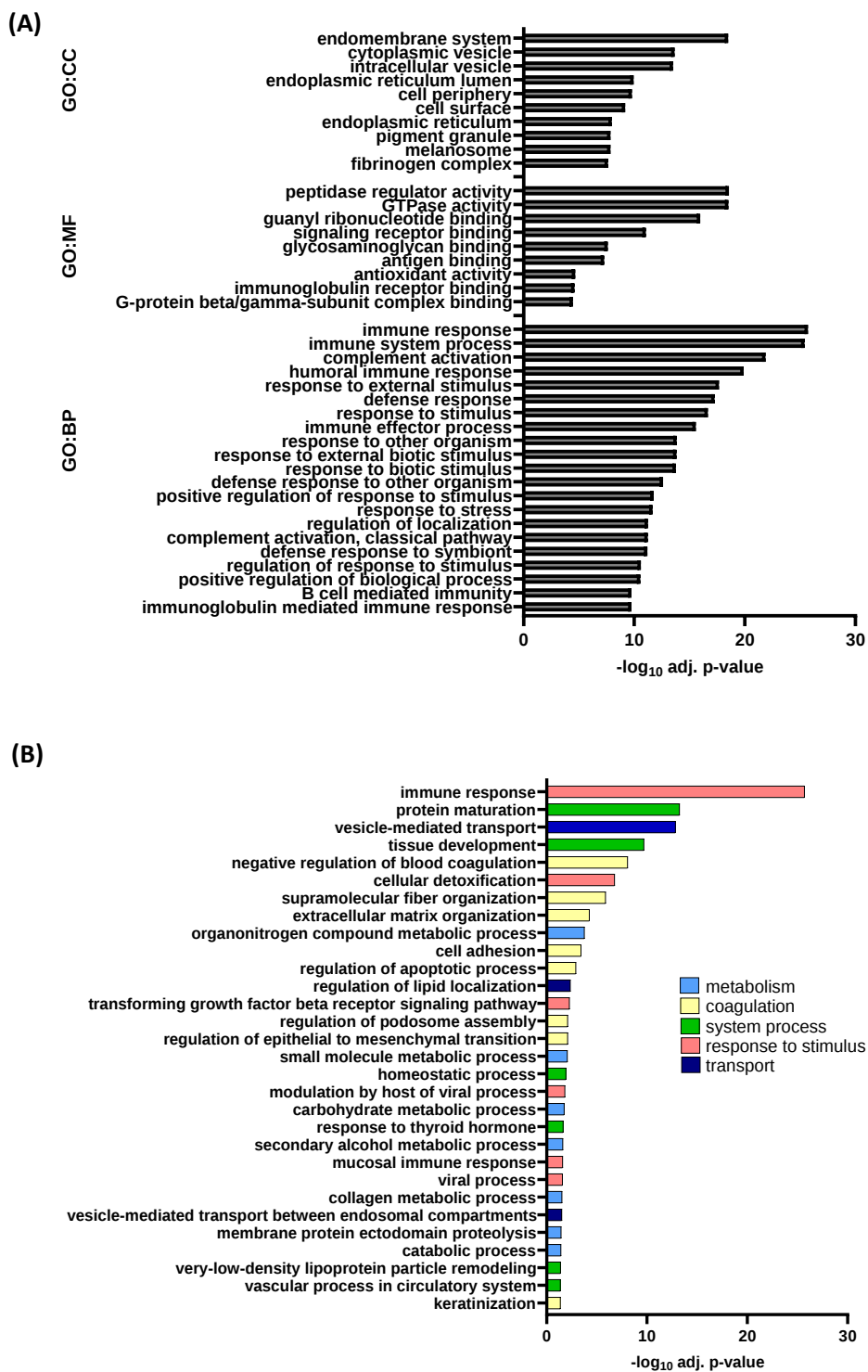


Figure 4.9. GO Enrichment Analysis of 634 quantified bovine milk proteins **(A)** with highlighted driver terms **(B)**.

CONCLUSION

Absolute concentrations of 634 bovine milk proteins were measured in bovine milk using LC-MS/MS bottom-up proteomics and BCA assays. Our study provided in-depth insights into the biological functions of bovine milk proteins and expanded the options for absolute quantification of proteins. This study bridges the gap between qualitative and quantitative bottom-up proteomics and paves the way for absolute quantification of every protein in food products. The proposed method is less resource-demanding than the conventional approaches, as it requires minimal proteins standards for absolute quantification of every protein.

REFERENCES

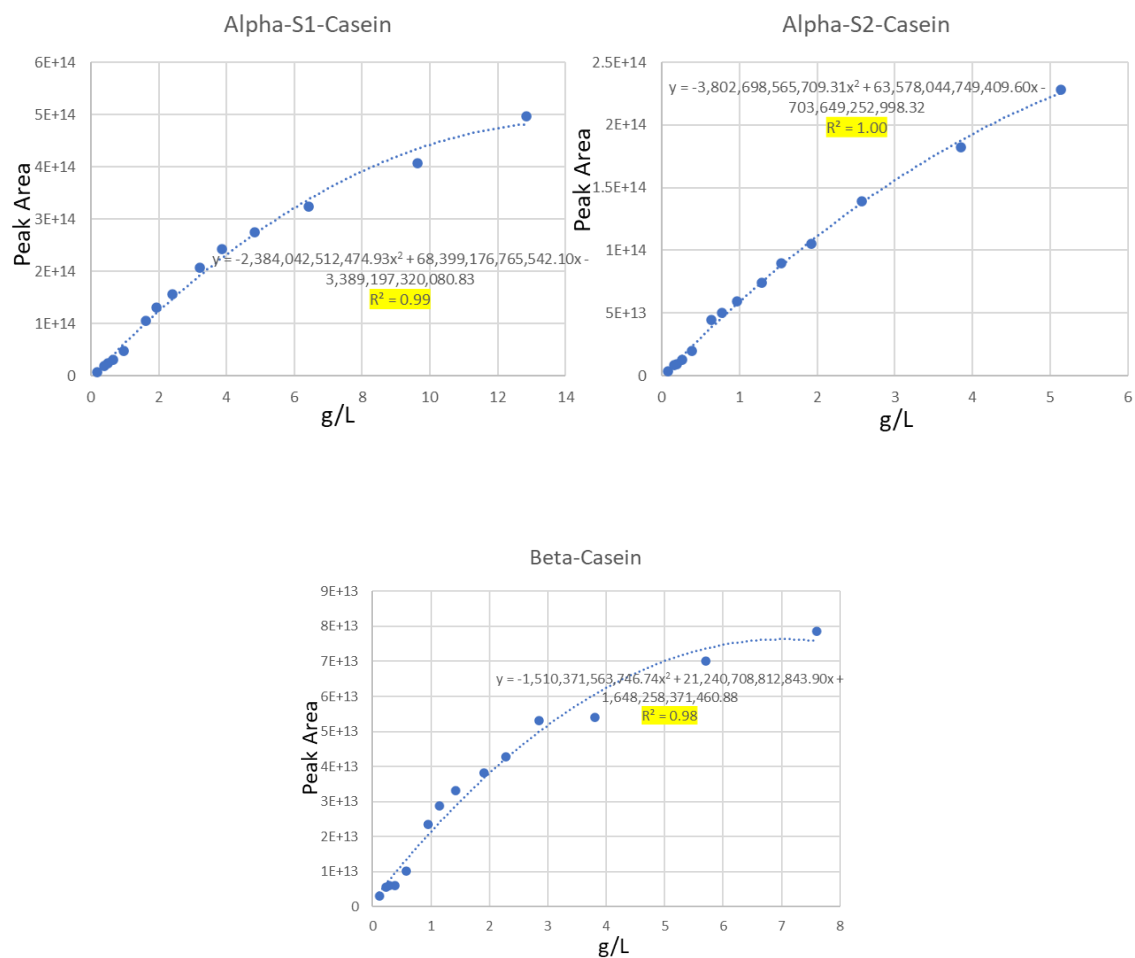
1. Hayes M. Measuring Protein Content in Food: An Overview of Methods. *Foods*. 2020 Sep 23;9(10):1340.
2. Mæhre H, Dalheim L, Edvinsen G, Elvevoll E, Jensen IJ. Protein Determination—Method Matters. *Foods*. 2018 Jan 1;7(1):5.
3. Langyan S, Bhardwaj R, Radhamani J, Yadav R, Gautam RK, Kalia S, et al. A Quick Analysis Method for Protein Quantification in Oilseed Crops: A Comparison With Standard Protocol. *Front Nutr*. 2022 May 31;9:892695.
4. Hueso D, Fontecha J, Gómez-Cortés P. Comparative study of the most commonly used methods for total protein determination in milk of different species and their ultrafiltration products. *Front Nutr*. 2022 Sep 13;9:925565.
5. Lynch JM, Barbano DM. Kjeldahl nitrogen analysis as a reference method for protein determination in dairy products. *J AOAC Int*. 1999;82(6):1389–98.
6. Chen Q, Ke X, Zhang JS, Lai SY, Fang F, Mo WM, et al. Proteomics method to quantify the percentage of cow, goat, and sheep milks in raw materials for dairy products. *Journal of Dairy Science*. 2016 Dec;99(12):9483–92.
7. Carrera M. Proteomics and Food Analysis: Principles, Techniques, and Applications. *Foods*. 2021 Oct 22;10(11):2538.
8. Stachniuk A, Sumara A, Montowska M, Fornal E. LIQUID CHROMATOGRAPHY–MASS SPECTROMETRY BOTTOM-UP PROTEOMIC METHODS IN ANIMAL SPECIES ANALYSIS OF PROCESSED MEAT FOR FOOD AUTHENTICATION AND THE DETECTION OF ADULTERATIONS. *Mass Spectrometry Reviews*. 2021 Jan;40(1):3–30.
9. Savitski MM, Mathieson T, Zinn N, Sweetman G, Doce C, Becher I, et al. Measuring and Managing Ratio Compression for Accurate iTRAQ/TMT Quantification. *J Proteome Res*. 2013 Aug 2;12(8):3586–98.
10. Pichler P, Köcher T, Holzmann J, Möhring T, Ammerer G, Mechtler K. Improved Precision of iTRAQ and TMT Quantification by an Axial Extraction Field in an Orbitrap HCD Cell. *Anal Chem*. 2011 Feb 15;83(4):1469–74.
11. Tacoma R, Fields J, Ebenstein DB, Lam YW, Greenwood SL. Characterization of the bovine milk proteome in early-lactation Holstein and Jersey breeds of dairy cows. *J Proteomics*. 2016 Jan 1;130:200–10.

12. Montowska M, Fornal E. Detection of peptide markers of soy, milk and egg white allergenic proteins in poultry products by LC-Q-TOF-MS/MS. *LWT*. 2018 Jan;87:310–7.
13. Goonatileke E, Huang J, Xu G, Wu L, Smilowitz JT, German JB, et al. Human Milk Proteins and Their Glycosylation Exhibit Quantitative Dynamic Variations during Lactation. *The Journal of Nutrition*. 2019 Aug;149(8):1317–25.
14. Huang J, Kailemia MJ, Goonatileke E, Parker EA, Hong Q, Sabia R, et al. Quantitation of human milk proteins and their glycoforms using multiple reaction monitoring (MRM). *Anal Bioanal Chem*. 2017 Jan;409(2):589–606.
15. Alvarez MRS, Zhou Q, Tena J, Lebrilla CB, Completo GC, Heralde FM, et al. *N*-Glycan and Glycopeptide Serum Biomarkers in Philippine Lung Cancer Patients Identified Using Liquid Chromatography–Tandem Mass Spectrometry. *ACS Omega*. 2022 Nov 8;7(44):40230–40.
16. Li Q, Xie Y, Wong M, Barboza M, Lebrilla CB. Comprehensive structural glycomic characterization of the glycocalyxes of cells and tissues. *Nat Protoc*. 2020 Aug;15(8):2668–704.
17. Bär C, Mathis D, Neuhaus P, Dürr D, Bisig W, Egger L, et al. Protein profile of dairy products: Simultaneous quantification of twenty bovine milk proteins. *International Dairy Journal*. 2019 Oct 1;97:167–75.
18. Maity S, Das D, Ambatipudi K. Quantitative alterations in bovine milk proteome from healthy, subclinical and clinical mastitis during *S. aureus* infection. *Journal of Proteomics*. 2020 Jul;223:103815.
19. Chatterton DEW, Aagaard S, Hesselballe Hansen T, Nguyen DN, De Gobba C, Lametsch R, et al. Bioactive proteins in bovine colostrum and effects of heating, drying and irradiation. *Food Funct*. 2020;11(3):2309–27.
20. Bai RJ, Li YS, Zhang FJ. Osteopontin, a bridge links osteoarthritis and osteoporosis. *Front Endocrinol*. 2022 Oct 28;13:1012508.
21. Cao X, Ren Y, Lu Q, Wang K, Wu Y, Wang Y, et al. Lactoferrin: A glycoprotein that plays an active role in human health. *Front Nutr*. 2023 Jan 5;9:1018336.
22. Koirala P, Dahal M, Rai S, Dhakal M, Nirmal NP, Maqsood S, et al. Dairy Milk Protein–Derived Bioactive Peptides: Avengers Against Metabolic Syndrome. *Curr Nutr Rep*. 2023 May 19;12(2):308–26.

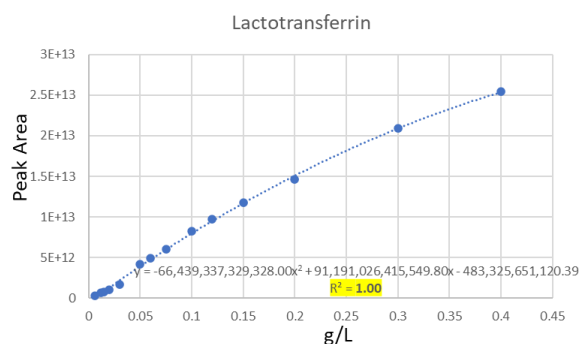
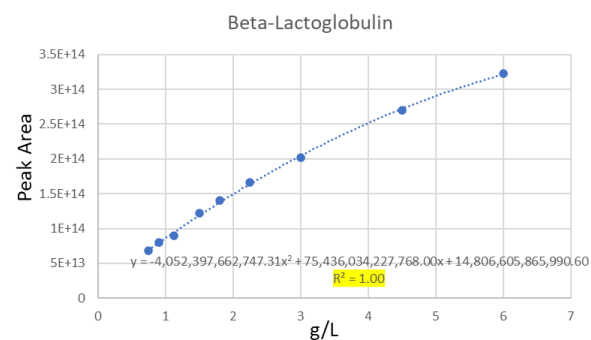
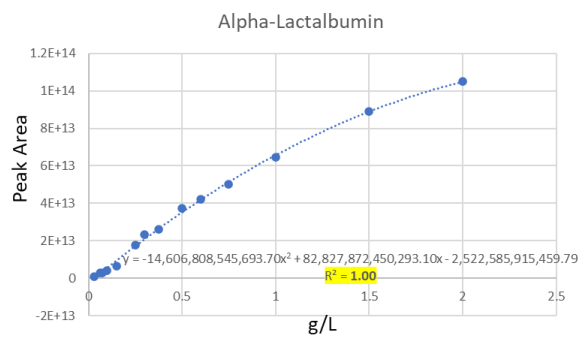
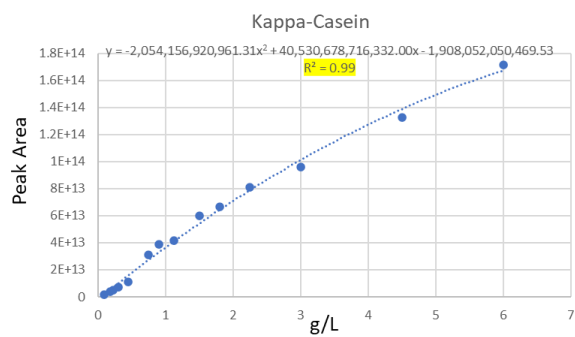
23. Bielecka M, Cichosz G, Cieczot H. Antioxidant, antimicrobial and anticarcinogenic activities of bovine milk proteins and their hydrolysates - A review. *International Dairy Journal*. 2022 Apr;127:105208.
24. Shimizu T, Ganzorig K, Miyamoto A, Ishii T, Urashima T, Fukuda K. A naturally occurring α_{s1} -casein-derived peptide in bovine milk inhibits apoptosis of granulosa cells induced by serum-free conditions: AN α_{s1} -CN-DERIVED PEPTIDE INHIBITS APOPTOSIS OF GRANULOSA CELLS. *J Pept Sci*. 2014 Mar;20(3):229–34.
25. Chatchatee P, Järvinen KM, Bardina L, Beyer K, Sampson HA. Identification of IgE- and IgG-binding epitopes on α_{s1} -casein: Differences in patients with persistent and transient cow's milk allergy. *Journal of Allergy and Clinical Immunology*. 2001 Feb;107(2):379–83.
26. Lacroix IME, Li-Chan ECY. Isolation and characterization of peptides with dipeptidyl peptidase-IV inhibitory activity from pepsin-treated bovine whey proteins. *Peptides*. 2014 Apr;54:39–48.
27. Bamdad F, Bark S, Kwon CH, Suh JW, Sunwoo H. Anti-Inflammatory and Antioxidant Properties of Peptides Released from β -Lactoglobulin by High Hydrostatic Pressure-Assisted Enzymatic Hydrolysis. *Molecules*. 2017 Jun 7;22(6):949.
28. Zhao Q, Wei G, Li K, Duan S, Ye R, Huang A. Identification and molecular docking of novel α -glucosidase inhibitory peptides from hydrolysates of Binglangjiang buffalo casein. *LWT*. 2022 Feb;156:113062.
29. Tu M, Wang C, Chen C, Zhang R, Liu H, Lu W, et al. Identification of a novel ACE-inhibitory peptide from casein and evaluation of the inhibitory mechanisms. *Food Chemistry*. 2018 Aug;256:98–104.
30. Tu M, Liu H, Zhang R, Chen H, Mao F, Cheng S, et al. Analysis and Evaluation of the Inhibitory Mechanism of a Novel Angiotensin-I-Converting Enzyme Inhibitory Peptide Derived from Casein Hydrolysate. *J Agric Food Chem*. 2018 Apr 25;66(16):4139–44.
31. He M, Sun J, Jiang ZQ, Yang YX. Effects of cow's milk beta-casein variants on symptoms of milk intolerance in Chinese adults: a multicentre, randomised controlled study. *Nutr J*. 2017 Dec;16(1):72.
32. Cui L, Yang G, Lu S, Zeng X, He J, Guo Y, et al. Antioxidant peptides derived from hydrolyzed milk proteins by *Lactobacillus* strains: A BIOPEP-UWM database-based analysis. *Food Research International*. 2022 Jun;156:111339.

33. Shazly AB, He Z, El-Aziz MA, Zeng M, Zhang S, Qin F, et al. Fractionation and identification of novel antioxidant peptides from buffalo and bovine casein hydrolysates. *Food Chemistry*. 2017 Oct;232:753–62.
34. Kim HJ, Yoon MJ, Lee J, Penninger JM, Kong YY. Osteoprotegerin Ligand Induces β -Casein Gene Expression through the Transcription Factor CCAAT/Enhancer-binding Protein β . *Journal of Biological Chemistry*. 2002 Feb;277(7):5339–44.
35. Chakrabarti S, Liao W, Davidge ST, Wu J. Milk-derived tripeptides IPP (Ile-Pro-Pro) and VPP (Val-Pro-Pro) differentially modulate angiotensin II effects on vascular smooth muscle cells. *Journal of Functional Foods*. 2017 Mar;30:151–8.
36. Rival SG, Boeriu CG, Wichers HJ. Caseins and Casein Hydrolysates. 2. Antioxidative Properties and Relevance to Lipoxygenase Inhibition. *J Agric Food Chem*. 2001 Jan 1;49(1):295–302.
37. Hubbard SR, Till JH. Protein Tyrosine Kinase Structure and Function. *Annu Rev Biochem*. 2000 Jun;69(1):373–98.
38. Huang Z, Zhang W, Gao C, Ji B, Chi X, Zheng W, et al. Dysregulation of GTPase IMAP family members in hepatocellular cancer. *Molecular Medicine Reports*. 2016 Nov;14(5):4119–23.
39. Korhonen HJ, Marnila P. Milk Bioactive Proteins and Peptides. In: Park YW, Haenlein GFW, editors. *Milk and Dairy Products in Human Nutrition* [Internet]. 1st ed. Wiley; 2013 [cited 2024 Feb 20]. p. 148–71. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/9781118534168.ch8>
40. Gaffen SL. Structure and signalling in the IL-17 receptor family. *Nat Rev Immunol*. 2009 Aug;9(8):556–67.

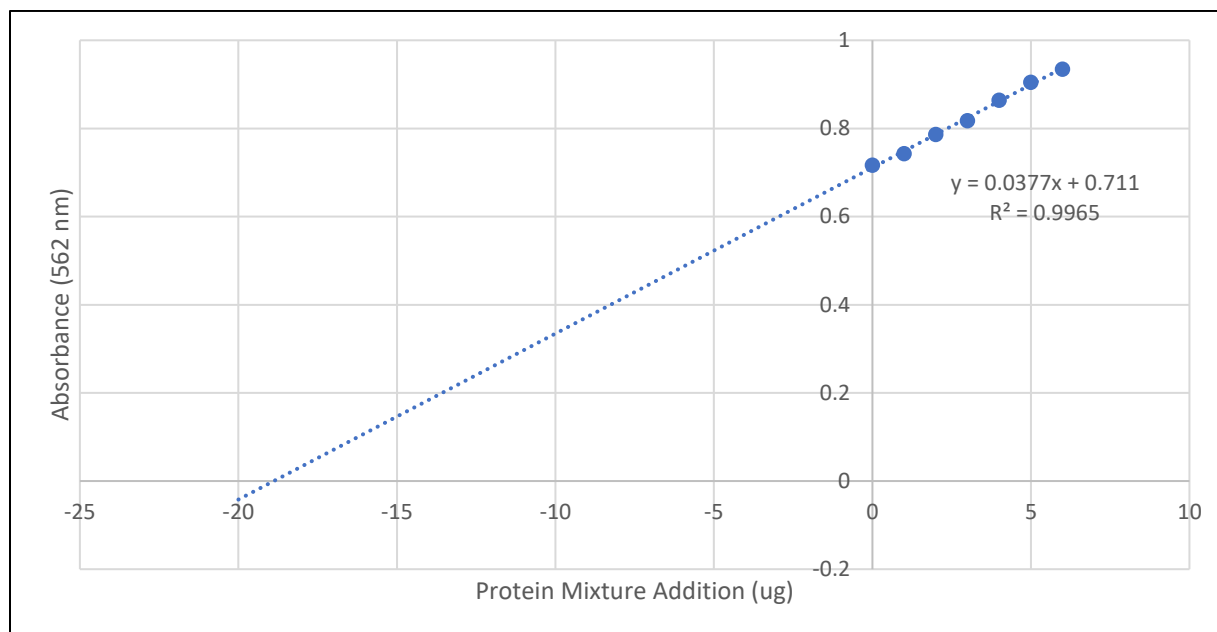
SUPPLEMENTARY INFORMATION



Supplementary Figure 4.1. Constructed calibration curves for quantification of bovine milk proteins based on peak areas of all peptides.



Supplementary Figure 4.1 Continued. Constructed calibration curves for quantification of bovine milk proteins based on peak areas of all peptides.



Supplementary Figure 4.2. Standard addition curve for absolute quantification of total milk protein using BCA.