

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Characterization of Glycocalyx for Biomarker Discovery in Cancer by Liquid Chromatography-Mass Spectrometry Methods

Permalink

<https://escholarship.org/uc/item/5j5989wm>

Author

Zhou, Qingwen

Publication Date

2021

Peer reviewed|Thesis/dissertation

Characterization of Glycocalyx for Biomarker Discovery in Cancer by Liquid Chromatography-Mass Spectrometry Methods

By

QINGWEN ZHOU
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Prof. Carlito B. Lebrilla, Chair

Prof. William Casey

Prof. Susan E. Ebeler

Committee in Charge

2021

DEDICATION

I want to dedicate this research to P.W. I hope my effort can help you accomplish your goals in life. I am sure great things will come about from your contribution to science.



*Qingwen Zhou
December 2014
Photo Credit: P.W.*

ACKNOWLEDGEMENTS

First, I would like to thank my parents (Guoqing Zhou and Yunping Chen). They moved to the United States 24 years ago to give me better opportunities in life. It has been a long journey for them, and now it is my turn to give back to them. I would also want to thank my grandparents, siblings, and nieces for their love and support throughout the year, especially my younger sister Julie.

I want to thank Dr. Carlito B. Lebrilla for giving me this opportunity to work on so many exciting projects. Additionally, I want to thank Dr. Carlito B. Lebrilla for the support, encouragement, and guidance throughout the years, especially during COVID, helping me out in a time of need. You are the best mentor anyone can ever ask for, both in science and in life. Thank you, Carlito!

I would like to give a special thanks to my cohort mates Juan Castillo, Ying Sheng, Jennyfer Tena, Christopher Ranque, AKA The Fab Five, for all the support throughout the year. We have all endured the same challenges, helping each other along the way. I will always cherish the time we had at UCD.

I would also like to thank the current members: Mariana Barboza, Ye Chen, Nikita Bacalzo, Garret Couture, Siyu Chen, Anita Vinjamuri, Ryan Schindler, Armin Oloumi, Cheng-Yu Weng, Yu Wang, Yasmine Bouchibti, Michael Russelle Alvarez, Aaron Stacy, and past members: Gege Xu, Diane Tu, Elisha Goonatilleke, Mathew Amicucci, Ace Galermo, Eshani Nandita, , Maurice Wong, Qiongyu Li, Selina Cheng, Yixuan Xie, Luster Mae N. Serrano, Kemal Solakyildirim, of the Lebrilla Lab You guys are like a second family to me and I wish the best for all of you.

I also want to give a special thanks to my undergraduates Cynthia Nguyen and Matthew Lam. You guys were terrific students, and the word brilliant would not even be able to describe how amazing you guys are. You guys have great potential in succeeding whatever comes your way.

I would like to also thank my dissertation committee members, Dr. William Casey and Dr. Susan E. Ebeler, for taking out time to revise my dissertation. I genuinely appreciate it. Thank you

I want to thank Brad Wolf for helping me start my Ph.D. and Emily Atkinson for helping me finish my Ph.D. I also want to thank Albert Sy and Anderson Ellis for all the technical support while at UCD, always offering great advice and great company.

I would also like to thank all mentors: Dr. Stefan Franzen, Dr. Kay Sandberg, Mrs. Crystal Molina, and Mrs. Ann S. McClung I had throughout the different stages of my life. I want to thank you for supporting me and believing in me. They were some of the best teachers anyone could ever ask for.

I want to give a special thanks to my best friends: Dr. Kevin Chang, Tyler Zhao, Yingcong Chen, and Alison Ke in life. You guys were here for all the rough times and supported me through them. Each one of you is brilliant in your own way. Thank you, my friends.

Lastly, I want to offer special thanks to my dear friend Sean William Houde and my loving grandpa Youshun Chen who passed away during my Ph.D. career. They both were amazing people who deserved more time to bring happiness to others, like how they brought joy in my life. I shall forever remember the time we spent together. May their soul rest in peace.

ABSTRACT

Glycosylation is one of the most prevalent post-translational modifications found on biomolecules that regulate various biological functions such as cell-cell interactions, immune responses, cellular regulations, viral/bacterial infections, and much more. Aberration in glycans can lead to alterations in protein function and have been shown to occur in many diseases such as cancer. In recent years, a few glycan-based biomarkers for cancer have been discovered. However, cancer is still a leading cause of death worldwide. Therefore, additional effort is needed to understand the cause of cancer and how glycosylation plays a role in cancer prognosis. This dissertation presents a series of state-of-the-art nano-liquid chromatography-mass spectrometry (nanoLC-MS) methods and bioinformatics tools to characterize the glycocalyx. In Chapter I, a thorough introduction of glycan synthesis, glycan biological functions, and methods of glycan analysis are presented. Chapter II presents the alteration of the N-glycome of several cancer cell lines through metabolic engineering using glycosylation inhibitors. Chapter III presents the first extensive “omic” analysis of the glycocalyx of 3D cell culture of cancer to elucidate alteration in the glycocalyx relative to its 2D counterpart. Finally, chapter IV presents the characterization of glycocalyx of lung cancer cell lines treated with a therapeutic drug to monitor glycocalyx alteration.

**THIS DISSERTATION IS BASED ON THE FOLLOWING
PUBLICATIONS:**

1. 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*, 2318.

ABBREVIATIONS

2FF: 2-deoxy-2-fluorofucose

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

6AF: 6-alkynyl fucose

C: Complex

CID: Collision Induced Dissociation

CV: Coefficient of Variation

ECC: Extracted Compound Chromatogram

EIC: Extracted Ion Chromatogram

ESI: Electrospray Ionization

FA: Formic Acid

Fuc: Fucose

Gal: Galactose

Glc: Glucose

GlcNAc: N-Acetyl-glucosamine

GTF: Glycosyltransferases

H: Hybrid

Hex: Hexose

HexNAc: N-Acetyl-hexosamine

HILIC: Hydrophilic Interaction Liquid Chromatography

HM: High-mannose

HPLC: High Performance Liquid Chromatography

Kif: Kifunensine

MALDI: Matrix-Assisted Laser Desorption Ionization

Man: Mannose

MS: Mass Spectrometry

Nano-LC: Nanoflow-Liquid Chromatography

NeuAc: N-Acetylneuraminic Acid

PGC: Porous Graphitized Carbon

PNGase F: Peptide N-Glycosidase F

PTM: Post-translational Modification

SPE: Solid-Phase Extraction

TIC: Total Ion Chromatogram

TOF: Time-of-Flight

UDP: Uridine Diphosphate

Table of Contents

Chapter I: Introduction to Structure, Biological Function, and Characterization of Glycans and Glycoproteins	1
Structure and Biosynthesis of Glycans and Glycoproteins	2
Biological Functions of Glycans	6
Method of Glycan Characterization	11
Conclusion.....	17
References	19
Chapter II: N-Glycomic Analysis of the Cell Shows Specific Effects of Glycosyl Transferase Inhibitors	23
Abstract	24
Introduction	24
Materials and Methods	27
Materials	27
Cell Line Culture	27
Cell Membrane Extraction.....	28
N-Glycan Release	29
NanoLC Chip–QTOF MS Analysis	29
Leica TCS SP8 STED 3X Fluorescence Imaging	30
Trypsin Digestion	30
Proteomic Data Analysis	31
Results	31
N-Glycan Profile of Cell Membranes.....	32
Cell Surface Proteomic with Inhibitors	39
Discussion	41
Conclusion.....	42

References	43
Supplementary Figures.....	46
Chapter III: Novel Characterization of the Glycocalyx of 3D Culturing Model Identified Glycans as Potential Biomarker for Colorectal Cancer	51
Abstract	52
Introduction	52
Methods and Materials	55
Cell Line Culture	55
Enzymatic Accutase Peeling of Spheroids	55
Cell Membrane Extraction.....	56
Protein Digestion	56
Glycocalyx Extraction	57
Glycocalyx nanoLC chip-QTOF Analysis	57
Proteomic nanoLC-MS/MS Analysis.....	58
Glycoproteomic nanoLC-MS/MS Analysis	59
Apoptosis Assay	59
Quantitative Proteomics and Glycoproteomics Data Analysis.....	60
Gene Ontology Analysis of Proteomics and Glycoproteomics	60
Glycoproteomic Profiling of ENPL and ITGA6	61
Results	61
N-Glycomic Profile of 2D and 3D Cell Membrane	63
Spatial N-Glycomic Profile of 3D Cell Membrane	66
Glycoproteomic Analysis of 2D and 3D Cell Membrane	71
Glycolipidomic Profile of 2D and 3D Cell Membrane	74
Proteomic analysis of 2D and 3D cell membrane	77

Discussions.....	80
Conclusion.....	83
References.....	84
Supplementary Figures.....	87
Chapter IV: Functional Characterization of an N-glycosylation Inhibitor Using Integrated Bioinformatics & Mass-spectrometric Glycomic-Driven Glycoproteomic Analysis	91
Abstract	92
Introduction	93
Materials and Methods	95
Network Pharmacology	95
In Silico Docking.....	96
Dose-Response Assay.....	97
Cell Cycle Assay	97
Apoptosis Assay	98
Scratch Assay	98
Trans-Well Migration Assay	98
Cell Treatment and Glycan, Protein, and Glycoprotein Enrichment.....	99
Glycomics Assay	99
Proteomics and Glycoproteomics Assay	100
Gene Ontology Analysis.....	102
Glycoprotein Molecular Modeling and Molecular Dynamics.....	102
Results	103
Pictilisib was Predicted to Interact with and Inhibit Glycosylation Enzymes Using In Silico Docking and Network Pharmacology.....	103
Pictilisib was Validated to Reduce the Relative Abundance of Fucosylated and Sialylated N-Glycans	105

Proteomic Analysis Shows Up-regulated Pathways Involving Apoptosis and Cell-Adhesion, and Down-Regulated Pathways Involving Cell Cycle Process, mRNA Processing, and Protein Translation	109
Glycoproteins with Reduced Fucosylation and Sialylation were Involved in Apoptosis, DNA Damage Repair, and Cell-Adhesion	113
Discussion	118
Conclusion.....	120
References	121
Supplementary Figures.....	127

Chapter I

Introduction to Structure, Biological Function, and Characterization of Glycans and Glycoproteins

oval circle represent monosaccharide, the green rectangle represents active sugar donor, and the asterisks are control points. Adapted from Freeze HH, Hart GW, Schnaar RL. Glycosylation Precursors. 2017. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 5. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK453043/> doi: 10.1101/glycobiology.3e.005

Monosaccharides are attached to glycans through the interconversion pathway, where they are converted to the active donor form. The active sugars are transferred to the glycans by glycosyltransferases (GTF) through the retaining or inverting mechanism²⁻⁴. Each active sugar has several different GTF, with each GTF transferring the monosaccharides to various positions and linkages on the glycans. Although we understand many aspects of glycans, such as how they are synthesized in the body or the linkage of which monosaccharides are orientated on glycans, there are still aspects that remain unknown, e.g., the incorporation efficiencies of different sugars.

Two of the most common glycans synthesized in humans and all organisms are N- and O-glycans. N-Glycans are bound to asparagine, and O-glycans are bound to serine or threonine residues. The glycans are synthesized and attached to proteins as part of the secretory pathway⁵. Unlike protein synthesis, which is template-driven, glycan synthesis is a non-template drive process with a series of, at times, competing enzymatic steps⁶.

N-Glycan synthesis begins with the dolichol pathway located in the endoplasmic reticulum (ER), where ALG7 (GlcNAc-1-phosphotransferase) transfers GlcNAc from uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) to dolichol phosphate to form N-acetyl- α -D-glucosaminyl-diphosphodolichol. This process starts a series of monosaccharides additions to

dolichol, forming the N-glycans precursor $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$. The synthetic pathway of the precursor to the final glycan product attached to a protein is shown in **Figure 1.2**⁷.

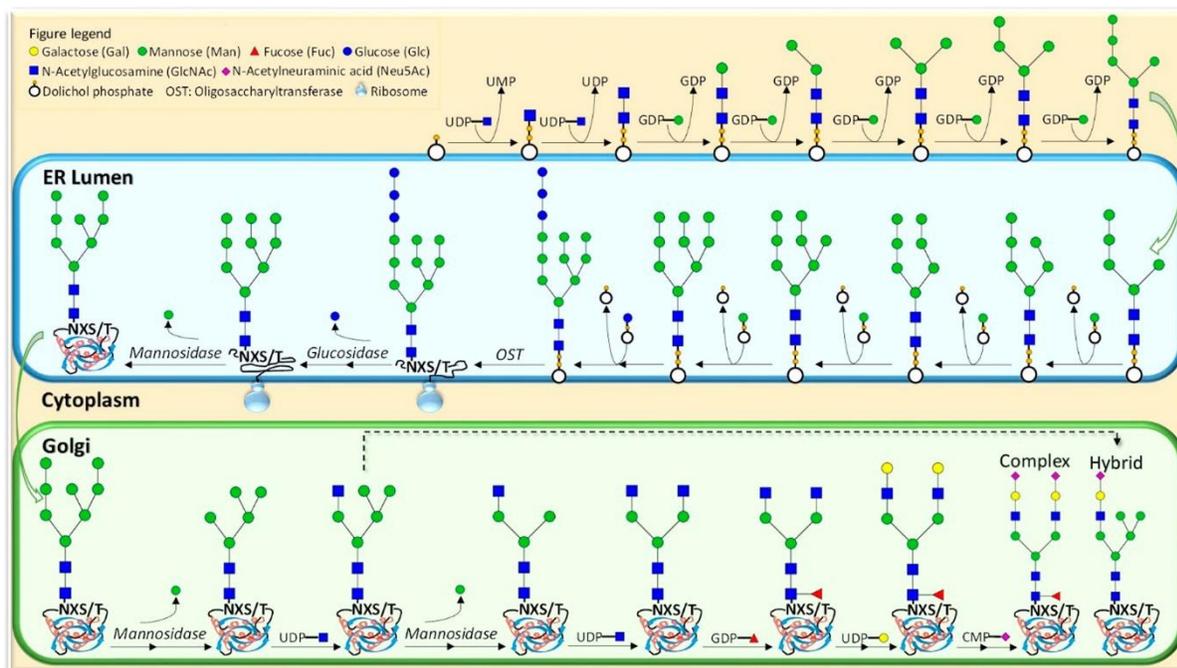


Figure 1.2 The N-glycan synthesis pathway begins with the dolichol pathway, where N-acetylglucosamine is attached to dolichol phosphate and ends in the Golgi. Within the Golgi, the N-glycans can be truncated, followed by decoration with other monosaccharides to form hybrid or complex type glycans. Reprinted with permission from Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatilleke, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chemical Reviews* 2018, 118 (17), 7886-7930. Copyright 2018 American Chemical Society.

$\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$ is then attached to an asparagine residue on a nascent polypeptide. Protein folding occurs once three glucose residues are removed from the glycan⁸. The glycoprotein is moved to the Golgi to allow additional truncation of the high-mannose structure, forming different composition variations such as fucosylated, sialylated, and

sialofucosylated glycans. O-Glycan synthesis does not require the initial lipid-linked oligosaccharide precursor to transfer glycans to proteins—the synthesis initiated by the attachment of GalNAc to a serine or threonine. GalNAc catalyzes the addition of other monosaccharides to form various compositions of O-glycans⁹. Both N-glycans and O-glycans have, respectively, common core structures. All N-glycans have a single core (Man α 1-3(Man α 1-6) Man β 1-4GlcNAc β 1-4GlcNAc β 1), which is bound directly to an asparagine residue with a consensus sequence of Asn-X-Ser/Thr (where X is any amino acid beside proline).

N-Glycans are categorized into three different types, namely high-mannose, hybrid, and complex-type N-glycans¹⁰. The basic structures three N-glycan types with the common core structure outlined in grey are displayed in **Figure 1.3**.

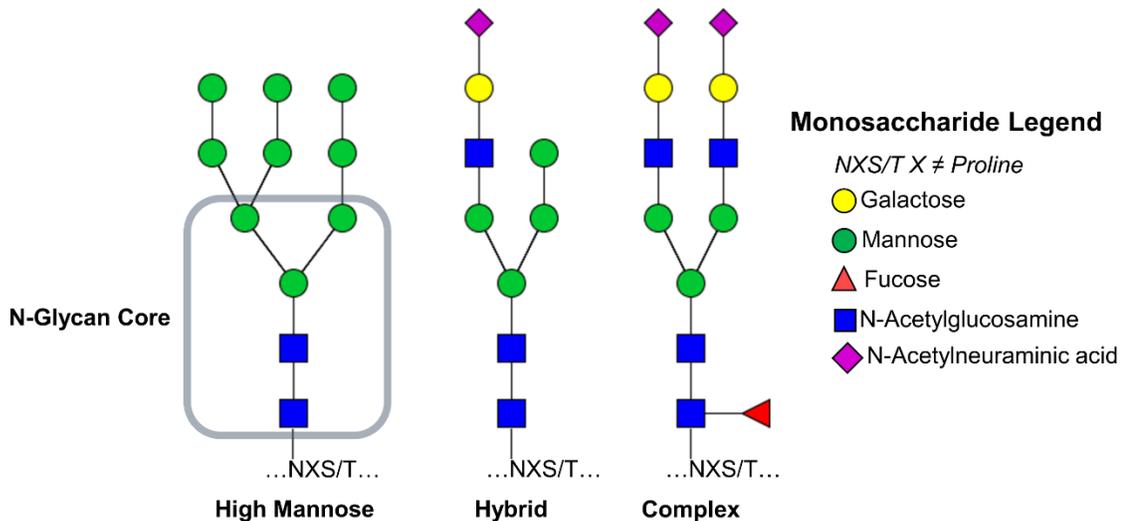


Figure 1.3 The three glycan types of N-glycans and the common core structures are highlighted in grey, covalently bound to an asparagine residue with a consensus sequence of Asn-X-Ser/Thr where X is any amino acid beside proline.

O-Glycans are attached to serine and threonine residues of the proteins but do not have a known consensus sequence. O-Glycans differ from N-glycans by the greater number of the core

structures in the former. O-Glycans have eight core structures, but only four cores are commonly seen. The common cores start with an α -linked GalNAc followed by the addition of Gal, GlcNAc, or both¹¹ and are displayed in **Figure 1.4**.

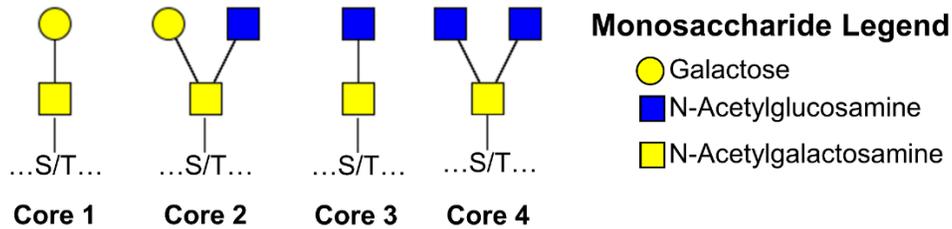


Figure 1.4 The four most common O-glycans structures are covalently bound to a serine or threonine residue without a consensus sequence.

Other core O-glycans can include other monosaccharides, such as α -linked O-fucose, β -linked O-GlcNAc, and α -linked O-mannose¹². N-Glycans and O-glycans are one of the most common post-translational modifications found on proteins. With the already large diversity in glycan structures, the further addition of glycans to proteins can add tremendous complexity and diversity to proteins.

Biological Functions of Glycans

Glycans play a vital role in many biological functions such as immune response, cell-to-cell interactions, and viral and bacterial protection (**Figure 1.5**)¹³. Their importance have led to extensive works to understand their functionality. However, many aspects of their functions are still unknown due to the difficulty of glycans analysis. One main challenge for understanding the functionality of glycans comes from the functional diversity of glycans. In addition, the structural complexity makes them challenging to characterize. However, the advancement in mass spectrometry methods, increased further our understanding of their functions. For example, it is known that O-fucose is vital for the Notch pathway, which is responsible for regulating cell

proliferation, cell fate, differentiation, and cell death^{14, 15}. In addition, studies have shown that the incorporation of fucose analogs caused steric clash with the Delta ligands of Notch EGG inhibiting Notch activity¹⁶. Additional functionalities of fucosylated glycans include as H-antigen in ABO blood groups, ligands for the selectin family of cell-adhesion receptors, and embryonic development¹⁷.

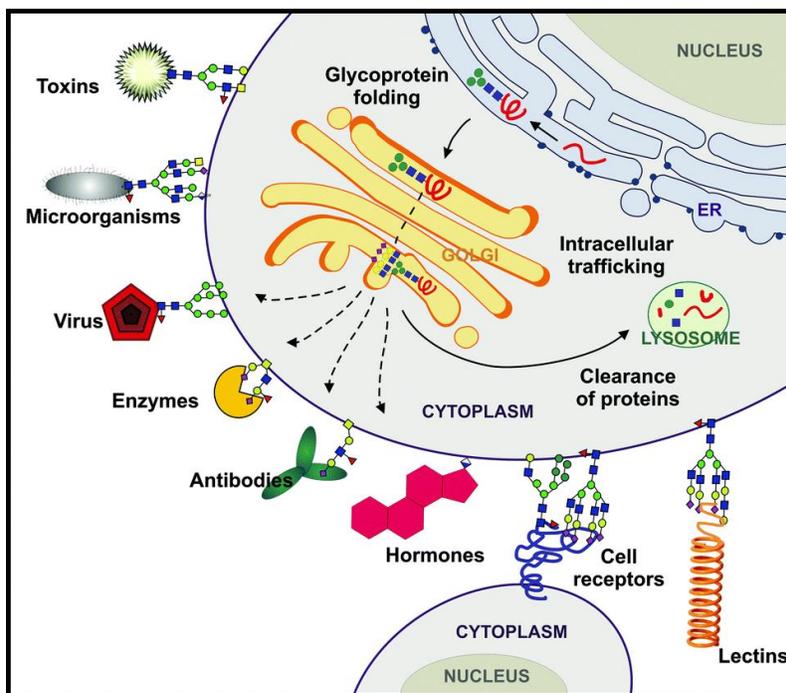


Figure 1.5 Examples of cellular functions regulated by glycans, namely, protein folding, host-microbe interactions, cell-cell communication, receptors for noxious agents, and receptor-ligand interactions. Reprint with permission from Defaus, S.; Gupta, P.; Andreu, D.; Gutiérrez-Gallego, R., Mammalian protein glycosylation--structure versus function. *Analyst* 2014, 139 (12), 2944-67. Copyright 2014 Royal Society of Chemistry.

Glycans also have protective functions in the immune system. Mucins are highly O-glycosylated glycoproteins known for their protective functions in our intestinal lining. They play a role as a physical barrier (chemical and biological barrier) and have antimicrobial

capabilities against infections¹⁸. There are several mucin glycoproteins ranging from MUC1 to MUC22 with each having its functional purpose¹⁹. Studies have shown that O-glycans are essential for mucin activities such as being physical barrier, shaping the intestinal ecosystem, and influencing the metabolic function of the microbiota²⁰.

As glycosylation is vital for maintaining homeostasis, the inability to metabolize glycoproteins due to genetic defects results in several congenital disorders^{21, 22}. For example, infants born with an α -Mannosidosis lysosomal disorder have progressive intellectual disability and only live between three to twelve years. On the other hand, if the α -Mannosidosis were instead to a β -Mannosidosis defect, the infant would likely have severe quadriplegia and death within 15 months^{23, 24}. These are only two among over 130 congenital disorders of glycosylation that are known^{25, 26}. In addition to glycosylation-related genetic defects, there are various diseases caused by aberrant glycosylation.

Aberrant glycosylation has been found in many deadly diseases, specifically cancer. Cancer is known for its abnormal proliferation rates with the ability to metastasize to other areas in the body²⁷. Numerous studies have shown that alteration in glycosylation is related to the onset of cancer, namely, an increase in core fucosylated glycans, high-mannose glycans, branching N-glycans, sialylated glycans, Thomsen-nouvelle antigen (Tn) and sialyl-Tn antigen (sTn). Examples of alteration in glycosylation are shown in **Figure 1.6**²⁸⁻³⁰.

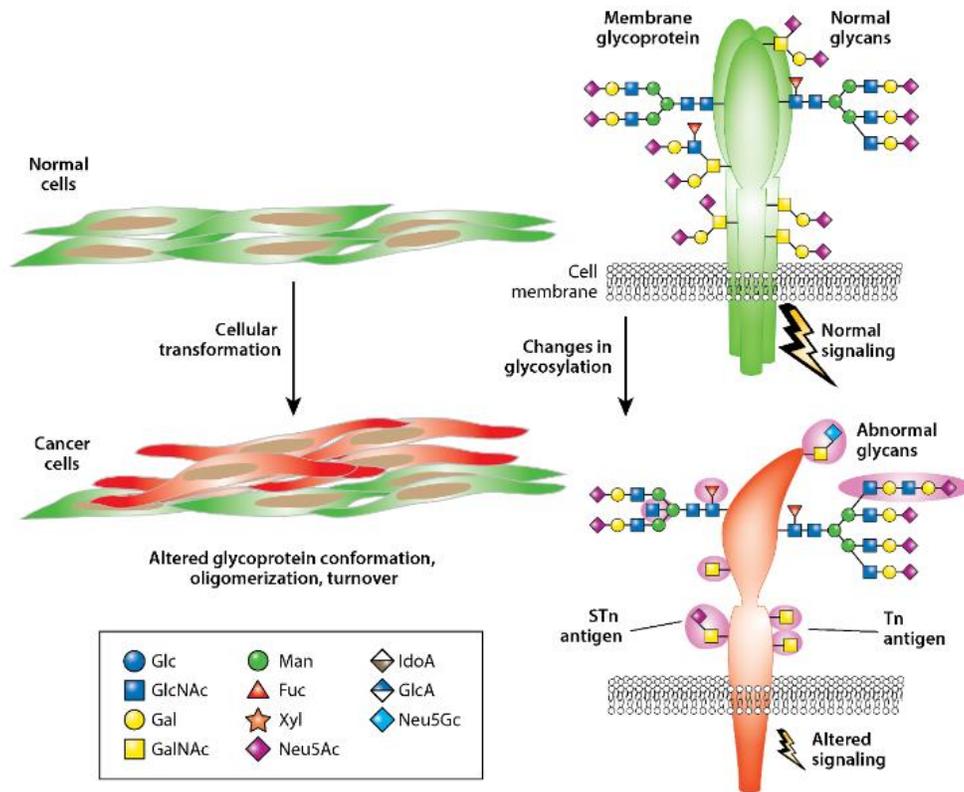


Figure 1.6 Common N- and O-glycosylation alteration found on cancerous cells include increases in core fucose, bisecting GluNAc, branching N-glycans, and STn and Tn antigens. Abbreviations: Fuc, fucose; Gal, galactose; GalNAc, N-acetylgalactosamine; Glc, glucose; GlcA, glucuronic acid; GlcNAc, N-acetylglucosamine; IdoA, iduronic acid; Man, mannose; Neu5Ac, 5-N-acetylneuraminic acid (sialic acid); Neu5Gc, 5-N-glycolylneuraminic acid; STn, sialyl-Tn; Xyl, xylose. Reprint with permission from Stowell, S. R.; Ju, T.; Cummings, R. D., Protein Glycosylation in Cancer. *Annual Review of Pathology: Mechanisms of Disease* 2015, 10 (1), 473-510. Copyright 2015 Annual Reviews.

Interestingly, most FDA-approved tumor markers are either glycoproteins or glycans. Some of the most common glycan's antigen, which are found on almost every cancer type, are Tn and sTn^{31,32}. Cancers that express a high level of sTn antigen are gastric, colorectal, pancreatic, cervical, endometrial, and ovarian³³. Another O-glycan that is known to be up-

regulated in tumor cells is O-GlcNAc. O-GlcNAc is known for regulating proteins involved in cell cycle progression, namely, FoxM1, cyclin D1, and cMYC. The upregulation of O-GlcNAcylation in cMYC protein competes with phosphorylation, stabilizing in cMYC protein, resulting in oncogenesis^{34, 35}.

Another glycan marker of cancer is the increased expression of sialic acid. As mentioned previously, an increase in sTn (sialylated O-glycan) antigen is a common trait of cancer progression. Other traits are increased expression in sialylated N-glycans and glycolipids (GD2). The upregulation of sialyltransferases causes increased sialylation. Sialyltransferases that are up-regulated in cancer include ST6GAL1, ST3GAL4, ST3GAL6, and ST6GalNAc1/2. These transferases are found in various types of cancer, including pancreatic, prostate, ovarian, and breast cancer. The increase in sialylated glycans promotes tumor growth, protecting cells from apoptosis, immune evasion, and facilitating cell detachment³⁶.

High-mannose type (specifically Man₉) N-glycans have been linked to increases in the progression of some cancers, most notably breast cancer³⁷. Modification of the N-glycome to express mainly high-mannose glycans using the inhibitor kifunensine has made cells more metastatic. Kifunensine is a MAN1A1 inhibitor that prevents the initial cleavage of alpha-1,2-linked mannose residues from Man₉GlcNAc₂. The metabolically engineered cells' wound-healing and migration rates were significantly increased compared to the control cell, suggesting the cells became more metastatic when expressing high-mannose glycans³⁸. Additionally, a therapeutic study conducted to target high-mannose glycans using lectin antibody conjugate (lectibody) showed potential therapeutic implications of blocking the activation of EGFR and IGFIR, which are known to be up-regulated in cancer³⁹. Although there is no universal trend in the N-glycan

profile during cancer progression, future studies will likely find more glycan-based biomarkers as we increase our understanding of glycosylation in cancer.

Method of Glycan Characterization

Glycans are among the most abundant post-translation modifications (PTM) found in proteins, and one of the most challenging to study due to the chemical and structural complexity. Because glycosylation of proteins plays a vital role in their biological functions, the ability to better characterize changes in glycan abundances, and structures associated with different diseases will increase our understanding the role of glycans in disease progression. The methods for analysis currently include lectin arrays, liquid chromatography-mass spectrometry (LC-MS), and matrix-assisted laser desorption/ionization - mass spectrometry (MALDI-MS) ⁴⁰⁻⁴².

Lectins-based methods such as lectin microarray and enzyme-linked lectin assays are already widely used to analyze glycosylation. Lectins are proteins (extracted mainly from plants) that specifically binds to carbohydrates. Common lectins include Concanavalin A (ConA), Elderberry lectin (SNA), and Aleuria aurantia lectin (AAL), each binding to a different carbohydrate. For example, AAL binds to core fucose, ConA binds to α -D-mannosyl, and SNA binds to sialic acid. In addition, lectins can also give linkage information. For example, SNA binds specifically to $\alpha(2,6)$ sialylated glycans. However, there are limitations with lectin binding assays. One is that the arrays only recognize motifs rather than entire structures, thus there could be little distinction between N-glycans, O-glycans, or glycolipids ⁴³⁻⁴⁵.

Matrix-assisted laser desorption/ionization – mass spectrometry (MALDI-MS) employs laser desorption techniques coupled to MS to analyze glycosylation. The advantage of MALDI-MS compared to other MS-based method is speed, it requires no separation, and it is amenable to

complicated mixtures. The released glycans are desorbed and ionized from the sample matrix by a laser pulse and then transferred into the mass analyzer (typically time-of-flight MS). The first glycan biomarker studies in serum for ovarian cancer were performed using MALDI-MS⁴⁶. However, MALDI-MS analysis of glycans has limitations such as lower sensitivity compared to other ionization methods techniques and can produce fragmentation of labile groups such as fucose and sialic acids⁴⁷. A more recent application of this technique for the analysis of oligosaccharides is in MALDI imaging. MALDI imaging MS provides spatial distributions of analytes on the sample surface. This method requires the detachment of glycans on tissue samples with the enzyme that releases N-glycans, peptide-N-glycosidase F (PNGase F). Further details of PNGase F will be discussed in a subsequent section. One capability of this method is that one can analyze the glycans of histopathological samples⁴⁸.

Liquid chromatography-mass spectrometry (LC-MS) is the more common MS-based method for analyzing glycans. These methods have been used to obtain structural information from N- and O-glycans and glycoconjugates such as glycolipids and glycoproteins. An enzymatic method using PNGase F is the most used method for N-glycans release. PNGase F is an amidase that cleaves the innermost GlcNAc breaking the covalent bond between the N-glycan and polypeptide. The glycoprotein mixture with PNGase F can be incubated overnight at 37 °C or microwaved to achieve complete release⁴⁹. PNGase F is currently the best method to release N-glycans because the enzyme can release any N-glycans except for those with $\alpha(1,3)$ -linked core fucose, which is not found in humans (**Figure 1.7**). Other glycosidases, such as PNGase A (cleave N-glycan with $\alpha(1,3)$ -linked core fucose), can release N-glycans at different positions; however, further details will not be covered in this work⁵⁰.

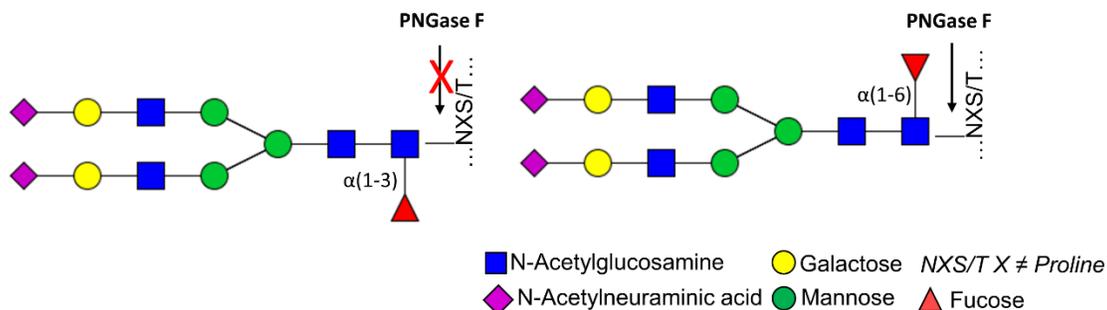


Figure 1.7 Peptide-N-glycosidase F (PNGase F) cleaves the innermost GlcNAc, breaking the covalent bond between the N-glycan and protein, but it cannot release any N-glycans with $\alpha(1,3)$ -linked core fucose.

There are commercially available O-glycosidases that can be used for O-glycans release, but none with broad substrate specificity like PNGase F. O-Glycans have eight different core structures, making it challenging to find an enzymatic method for O-glycan release. Therefore, chemical release is required for the complete release of O-glycans. The most commonly used approach is reductive β -elimination under alkaline conditions shown in **Figure 1.8**⁵¹. Additional methods of O-glycan release can be performed with ordinary household bleach or hydrazinolysis^{52 53}. A limitation of the chemical releasing O-glycans is the poor reproducibility. Peeling of the released glycans usually occurs during chemical release because of the harsh conditions required for the reactions⁵¹. After the glycans are released, glycomic analysis can be performed on the free N- and O-glycans using LC-MS.

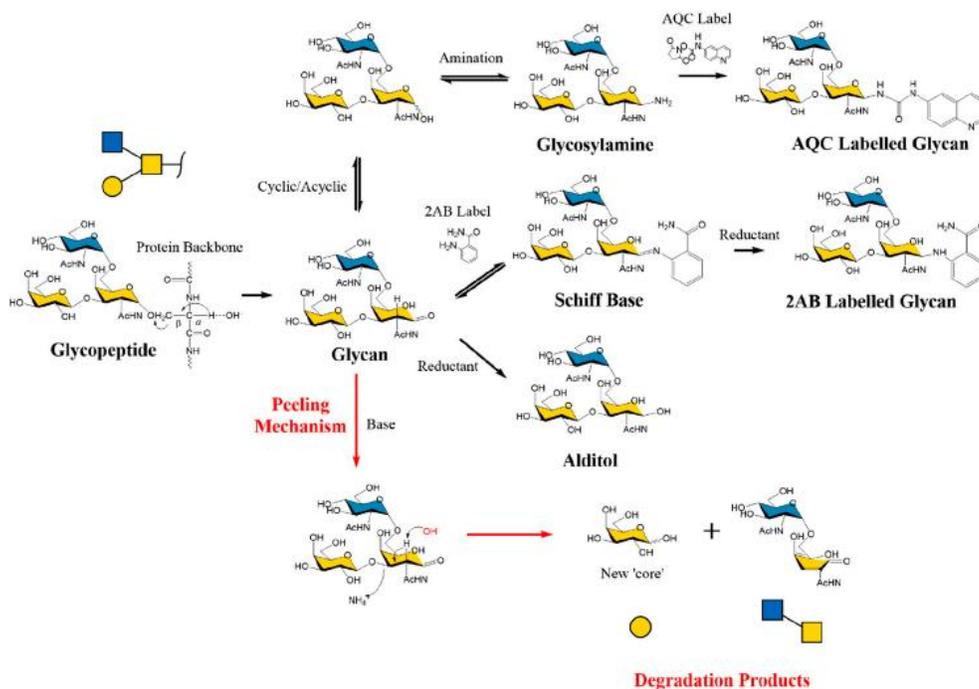


Figure 1.8 Methods of O-glycan release such as β -elimination, peeling mechanism, and end-capping strategies include label-free and label methods of analysis. Reprinted with permission from Wilkinson, H.; Saldova, R., Current Methods for the Characterization of O-Glycans.

Journal of Proteome Research 2020, 19 (10), 3890-3905. Copyright 2020 American Chemical Society.

Currently, the most prevalent method for glycan analysis is using an electrospray ionization (ESI) source coupled with an LC-MS instrument. ESI is widely used because this ionization method is a soft ionization process, reducing in-source fragmentation. On the other hand, LC-MS methods have a higher level sensitivity for glycans analysis, especially when coupling with a nanoLC system. In addition, liquid chromatography has various separation techniques such as reverse-phase liquid chromatography (RPLC), hydrophilic interaction chromatography (HILIC), and porous graphitized carbon (PGC) that are beneficial for glycan analysis. PGC is the most commonly used stationary phase for glycan separation and purification

since it is the most effective technique at separating native glycans⁵⁴. An N-glycan analysis using a nanoLC-MS with PGC stationary phase of a cell line is shown in **Figure 1.9**. Although, PGC has limitations in extracting linkage information.

On the other hand, RPLC only works for permethylated and derivatized glycans since native glycans are hydrophilic and cannot retain on the C18 stationary phase. With this approach, linkage information can be obtained by analyzing permethylated and derivatized glycans data⁵⁵. However, reverse-phase C18 is more commonly used for peptide and glycopeptide analysis. Further discussion will be covered in subsequent sections.

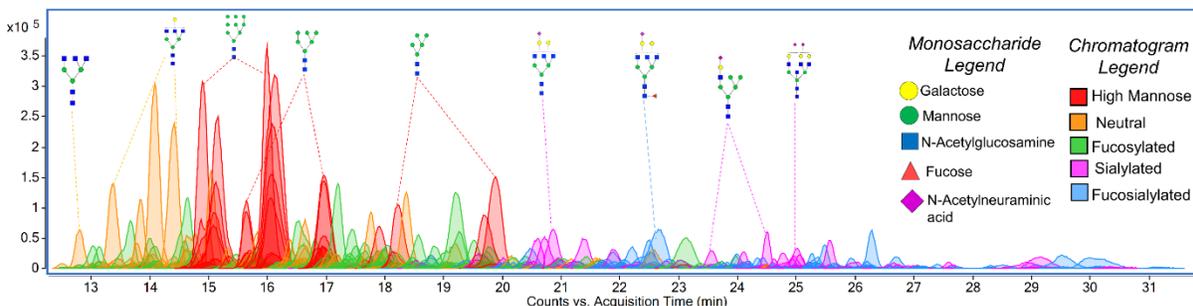


Figure 1.9 The N-glycan profile of HT29 shows glycans' separation and identification (with isomers) using a nanoLC-MS instrument with a PGC stationary phase

After the glycans are separated in the LC, they are injected into the MS instrument. Untargeted glycan analysis uses a QTOF mass spectrometry, while targeted analysis uses a triple quadrupole (QqQ) mass spectrometry. Using MS to analyze glycan enables for femto- to attomolar level of detection, allowing the analysis of small amounts of materials. This is advantageous for glycan analysis since glycan sample sizes are typically limited and the analytes are in relatively low abundance. Newer methods of glycan analysis used an Orbitrap MS, which has a higher level of sensitivity than other MS analyzers⁵⁶. However, Orbitrap MS is usually used for glycopeptide analysis and will be discussed in subsequent sections. Another benefit of

MS analysis is to obtain structural information from tandem MS data through collision-induced dissociation (CID) fragmentation -the most widely used fragmentation method for glycans.

Figure 1.10 depicts the primary cleavage site for native glycan fragmentation.

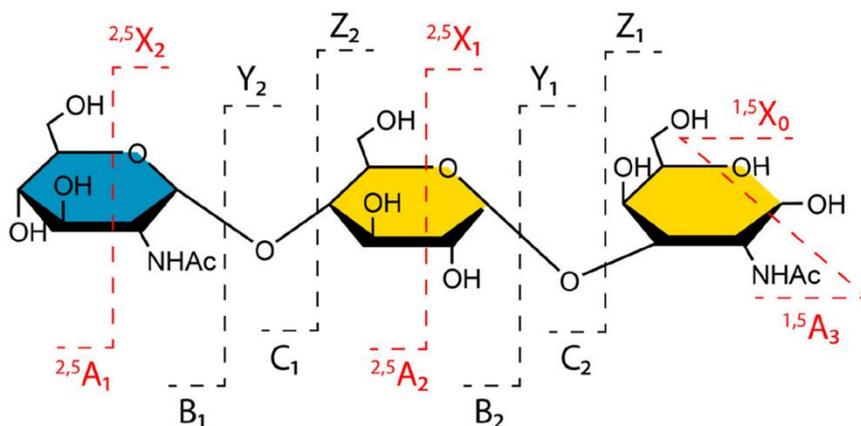


Figure 1.10 Fragmentation pattern of a glycan under collision-induced dissociation. Reprint with permission from Wilkinson, H.; Saldova, R., Current Methods for the Characterization of O-Glycans. *Journal of Proteome Research* 2020, 19 (10), 3890-3905. Copyright 2020 American Chemical Society.

Current glycoprotein analysis uses trypsin digestion and HILIC enrichment for the untargeted glycopeptide analysis. Standard glycoprotein procedures start with protein denaturation with DTT and heat followed by alkylation to inhibit the reformation of the disulfide bond. Thereafter, the proteins are digested with trypsin at 37 °C for 18 hours, and the glycopeptides are enriched with HILIC solid-phase extraction (SPE). The enrichment process is necessary to reduce the signal suppression from the nonglycosylated peptide. Finally, the samples are separated in a C18 nano column prior to injection to the Orbitrap MS. Orbitrap instruments are the best analyzer to perform untargeted analysis because of the high resolution

and mass accuracy. For glycopeptide fragmentation, step higher-energy collisional dissociation (HCD) was used to produce fragmentation of both the peptide backbone and attached glycans⁵⁷.

Targeted glycopeptide analysis (using MRM) includes ultra-performance liquid chromatography (UPLC) for separation. Targeted methods select specific ionic species, which remove the need for glycopeptide enrichment. Samples are still digested with trypsin. However, a study has shown that the addition of serine protease Glu-C with trypsin can increase the number of glycopeptides identified⁵⁸. Both methods use C18 as the stationary phase in the chromatographic separation. The C18 stationary phase mainly binds with the peptide backbone and retention of the glycopeptide can be predicted based on the amino acid sequence⁵⁹. Currently, there are only a few software programs that can do glycopeptide analysis, such as Byonic and pGlyco. Additional software development would be required to support the need of the field sufficiently.

Conclusion

Glycosylation is one of the most prevalent types of Post Translational Modification found in organisms. Glycans play a vital role in many biological functions, including immune response, cell-cell interactions, and cell regulation. New analytical methods are constantly being developed to advance glycoalyx analysis. In the following chapters, a novel platform for glycoalyx analysis by LC-MS will be introduced. The presented work combines multiple “omic” analyses and bioinformatic tools to identify and profile alterations caused by glycosylation inhibitors, cell culturing methods, and therapeutic drugs on various cancer cell lines. Chapter 2 will present the effect of glycosylation inhibitors on various cancer cell lines. Chapter 3 will present the difference between the 2D and 3D culturing methods and the effect of protein and glycan

expression on the glycocalyx. Finally, Chapter 4 will investigate cancer therapeutics and their effects on glycosylation.

References

1. Freeze, H. H.; Hart, G. W.; Schnaar, R. L., Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2017.
2. Schuman, B.; Evans, S. V.; Fyles, T. M., Geometric attributes of retaining glycosyltransferase enzymes favor an orthogonal mechanism. *PLoS One* 2013, 8 (8), e71077.
3. Kozmon, S.; Tvaroška, I., Catalytic Mechanism of Glycosyltransferases: Hybrid Quantum Mechanical/Molecular Mechanical Study of the Inverting N-Acetylglucosaminyltransferase I. *Journal of the American Chemical Society* 2006, 128 (51), 16921-16927.
4. Lee, S. S.; Hong, S. Y.; Errey, J. C.; Izumi, A.; Davies, G. J.; Davis, B. G., Mechanistic Evidence for a Front-Side, SNi-type Reaction in a Retaining Glycosyltransferase. *Nat Chem Biol* 2011, 7 (9), 631-8.
5. Adams, B. M.; Canniff, N. P.; Guay, K. P.; Larsen, I. S. B.; Hebert, D. N., Quantitative Glycoproteomics Reveals Cellular Substrate Selectivity of the ER Protein Quality Control Sensors UGGT1 and UGGT2. *Elife* 2020, 9.
6. Lopez Aguilar, A.; Briard, J. G.; Yang, L.; Ovrzyn, B.; Macauley, M. S.; Wu, P., Tools for Studying Glycans: Recent Advances in Chemoenzymatic Glycan Labeling. *ACS Chemical Biology* 2017, 12 (3), 611-621.
7. Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatileke, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chemical Reviews* 2018, 118 (17), 7886-7930.
8. Shental-Bechor, D.; Levy, Y., Effect of Glycosylation on Protein Folding: A Close Look at Thermodynamic Stabilization. *Proceedings of the National Academy of Sciences* 2008, 105 (24), 8256-8261.
9. Wopereis, S.; Lefeber, D. J.; Morava, E.; Wevers, R. A., Mechanisms in Protein O-Glycan Biosynthesis and Clinical and Molecular Aspects of Protein O-Glycan Biosynthesis Defects: A Review. *Clin Chem* 2006, 52 (4), 574-600.
10. 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 Mar;100:94-100.
11. Sjögren, J.; Collin, M., Bacterial Glycosidases in Pathogenesis and Glycoengineering. *Future Microbiology* 2014, 9, 1039-51.
12. You, X.; Qin, H.; Ye, M., Recent Advances in Methods for the Analysis of Protein O-Glycosylation at Proteome Level. *J Sep Sci* 2018, 41 (1), 248-261.
13. Defaus, S.; Gupta, P.; Andreu, D.; Gutiérrez-Gallego, R., Mammalian Protein Glycosylation-Structure Versus Function. *Analyst* 2014, 139 (12), 2944-67.
14. Brückner, K.; Perez, L.; Clausen, H.; Cohen, S., Glycosyltransferase Activity of Fringe Modulates Notch–Delta Interactions. *Nature* 2000, 406 (6794), 411-415.
15. Kopan, R., Notch signaling. *Cold Spring Harb Perspect Biol* 2012, 4 (10), a011213.
16. Schneider, M.; Kumar, V.; Nordstrøm, L. U.; Feng, L.; Takeuchi, H.; Hao, H.; Luca, V. C.; Garcia, K. C.; Stanley, P.; Wu, P.; Haltiwanger, R. S., Inhibition of Delta-induced Notch Signaling Using Fucose Analogs. *Nature Chemical Biology* 2018, 14 (1), 65-71.
17. Becker, D. J.; Lowe, J. B., Fucose: Biosynthesis and Biological Function in Mammals. *Glycobiology* 2003, 13 (7), 41R-53R.
18. Linden, S. K.; Sutton, P.; Karlsson, N. G.; Korolik, V.; McGuckin, M. A., Mucins in the Mucosal Barrier to Infection. *Mucosal Immunology* 2008, 1 (3), 183-197.

19. Moniaux, N.; Escande, F.; Porchet, N.; Aubert, J. P.; Batra, S. K., Structural Organization and Classification of the Human Mucin Genes. *Front Biosci* 2001, 6, D1192-206.
20. Bergstrom, K. S. B.; Xia, L., Mucin-Type O-Glycans and Their Roles in Intestinal Homeostasis. *Glycobiology* 2013, 23 (9), 1026-1037.
21. Kudelka, M. R.; Stowell, S. R.; Cummings, R. D.; Neish, A. S., Intestinal Epithelial Glycosylation in Homeostasis and Gut Microbiota Interactions in IBD. *Nature Reviews Gastroenterology & Hepatology* 2020, 17 (10), 597-617.
22. E F Neufeld; T W Lim, a.; Shapiro, L. J., Inherited Disorders of Lysosomal Metabolism. *Annual Review of Biochemistry* 1975, 44 (1), 357-376.
23. Beck, M.; Olsen, K. J.; Wraith, J. E.; Zeman, J.; Michalski, J.-C.; Saftig, P.; Fogh, J.; Malm, D., Natural History of Alpha Mannosidosis a Longitudinal Study. *Orphanet J Rare Dis* 2013, 8, 88-88.
24. Neufeld, E. F.; Lim, T. W.; Shapiro, L. J., Inherited Disorders of Lysosomal Metabolism. *Annu Rev Biochem* 1975, 44, 357-76.
25. Péanne, R.; de Lonlay, P.; Foulquier, F.; Kornak, U.; Lefeber, D. J.; Morava, E.; Pérez, B.; Seta, N.; Thiel, C.; Van Schaftingen, E.; Matthijs, G.; Jaeken, J., Congenital Disorders of Glycosylation (CDG): Quo vadis? *Eur J Med Genet* 2018, 61 (11), 643-663.
26. Ng, B. G.; Freeze, H. H., Human Genetic Disorders Involving Glycosylphosphatidylinositol (GPI) Anchors and Glycosphingolipids (GSL). *J Inherit Metab Dis* 2015, 38 (1), 171-178.
27. Munkley, J.; Elliott, D. J., Hallmarks of Glycosylation in Cancer. *Oncotarget* 2016, 7 (23), 35478-89.
28. Stowell, S. R.; Ju, T.; Cummings, R. D., Protein Glycosylation in Cancer. *Annual Review of Pathology: Mechanisms of Disease* 2015, 10 (1), 473-510.
29. Osuga, T.; Takimoto, R.; Ono, M.; Hirakawa, M.; Yoshida, M.; Okagawa, Y.; Uemura, N.; Arihara, Y.; Sato, Y.; Tamura, F.; Sato, T.; Iyama, S.; Miyanishi, K.; Takada, K.; Hayashi, T.; Kobune, M.; Kato, J., Relationship Between Increased Fucosylation and Metastatic Potential in Colorectal Cancer. *J Natl Cancer Inst* 2016, 108 (9): djw038.
30. Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J., Glycosylation in Health and Disease. *Nat Rev Nephrol* 2019, 15 (6), 346-366.
31. Hammarström, S., The Carcinoembryonic Antigen (CEA) Family: Structures, Suggested Functions and Expression in Normal and Malignant Tissues. *Seminars in Cancer Biology* 1999, 9 (2), 67-81.
32. Häuselmann, I.; Borsig, L., Altered Tumor-Cell Glycosylation Promotes Metastasis. *Front Oncol* 2014, 4, 28-54.
33. Munkley, J., The Role of Sialyl-Tn in Cancer. *Int J Mol Sci* 2016, 17 (3), 275.
34. Itkonen, H. M.; Minner, S.; Guldvik, I. J.; Sandmann, M. J.; Tsourlakis, M. C.; Berge, V.; Svindland, A.; Schlomm, T.; Mills, I. G., O-GlcNAc Transferase Integrates Metabolic Pathways to Regulate the Stability of c-MYC in Human Prostate Cancer Cells. *Cancer Res* 2013, 73 (16), 5277-87.
35. Caldwell, S. A.; Jackson, S. R.; Shahriari, K. S.; Lynch, T. P.; Sethi, G.; Walker, S.; Vosseller, K.; Reginato, M. J., Nutrient Sensor O-GlcNAc Transferase Regulates Breast Cancer Tumorigenesis Through Targeting of the Oncogenic Transcription Factor FoxM1. *Oncogene* 2010, 29 (19), 2831-42.
36. Munkley, J.; Scott, E., Targeting Aberrant Sialylation to Treat Cancer. *Medicines (Basel)* 2019, 6 (4), 102.

37. de Leoz, M. L.; Young, L. J.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B., High-mannose Glycans are Elevated During Breast Cancer Progression. *Mol Cell Proteomics* 2011, 10 (1), M110.002717.
38. Park, D. D.; Phoomak, C.; Xu, G.; Olney, L. P.; Tran, K. A.; Park, S. S.; Haigh, N. E.; Luxardi, G.; Lert-Itthiporn, W.; Shimoda, M.; Li, Q.; Matoba, N.; Fierro, F.; Wongkham, S.; Maverakis, E.; Lebrilla, C. B., Metastasis of Cholangiocarcinoma is Promoted by Extended High-Mannose Glycans. *Proc Natl Acad Sci U S A* 2020, 117 (14), 7633-7644.
39. Oh, Y. J.; Dent, M. W.; Freels, A. R.; Zhou, Q.; Lebrilla, C. B.; Merchant, M. L.; Matoba, N., Antitumor Activity of a Lectin Targeting Cancer-Associated High-Mannose Glycans. *bioRxiv* 2021, 2021.04.28.441869.
40. Geissner, A.; Reinhardt, A.; Rademacher, C.; Johannessen, T.; Monteiro, J.; Lepenies, B.; Thépaut, M.; Fieschi, F.; Mrázková, J.; Wimmerova, M.; Schuhmacher, F.; Götze, S.; Grünstein, D.; Guo, X.; Hahn, H. S.; Kandasamy, J.; Leonori, D.; Martin, C. E.; Parameswarappa, S. G.; Pasari, S.; Schlegel, M. K.; Tanaka, H.; Xiao, G.; Yang, Y.; Pereira, C. L.; Anish, C.; Seeberger, P. H., Microbe-Focused Glycan Array Screening Platform. *Proceedings of the National Academy of Sciences* 2019, 116 (6), 1958-1967.
41. 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.
42. Han, L.; Costello, C. E., Mass Spectrometry of Glycans. *Biochemistry (Mosc)* 2013, 78 (7), 710-720.
43. Rillahan, C. D.; Paulson, J. C., Glycan Microarrays for Decoding the Glycome. *Annual Review of Biochemistry* 2011, 80, 797-823.
44. Ghazarian, H.; Itoni, B.; Oppenheimer, S. B., A Glycobiology Review: Carbohydrates, Lectins and Implications in Cancer Therapeutics. *Acta Histochem* 2011, 113 (3), 236-247.
45. Xie, Y.; Sheng, Y.; Li, Q.; Ju, S.; Reyes, J.; Lebrilla, C. B., Determination of the Glycoprotein Specificity of Lectins on Cell Membranes Through Oxidative Proteomics. *Chemical Science* 2020, 11 (35), 9501-9512.
46. An, H. J.; Miyamoto, S.; Lancaster, K. S.; Kirmiz, C.; Li, B.; Lam, K. S.; Leiserowitz, G. S.; Lebrilla, C. B., Profiling of Glycans in Serum for the Discovery of Potential Biomarkers for Ovarian Cancer. *J Proteome Res* 2006, 5 (7), 1626-35.
47. Sekiya, S.; Wada, Y.; Tanaka, K., Derivatization for Stabilizing Sialic Acids in MALDI-MS. *Analytical Chemistry* 2005, 77 (15), 4962-4968.
48. Powers, T. W.; Jones, E. E.; Betesh, L. R.; Romano, P. R.; Gao, P.; Copland, J. A.; Mehta, A. S.; Drake, R. R., Matrix Assisted Laser Desorption Ionization Imaging Mass Spectrometry Workflow for Spatial Profiling Analysis of N-Linked Glycan Expression in Tissues. *Analytical Chemistry* 2013, 85 (20), 9799-9806.
49. Sandoval, W. N.; Arellano, F.; Arnott, D.; Raab, H.; Vandlen, R.; Lill, J. R., Rapid Removal of N-Linked Oligosaccharides Using Microwave Assisted Enzyme Catalyzed Deglycosylation. *International Journal of Mass Spectrometry* 2007, 259 (1), 117-123.
50. TRETTER, V.; ALTMANN, F.; MÄRZ, L., Peptide-N4-(N-acetyl- β -glucosaminyl)asparagine Amidase F Cannot Release Glycans with Fucose Attached $\alpha 1 \rightarrow 3$ to the Asparagine-Linked N-Acetylglucosamine Residue. *European Journal of Biochemistry* 1991, 199 (3), 647-652.
51. Wilkinson, H.; Saldova, R., Current Methods for the Characterization of O-Glycans. *Journal of Proteome Research* 2020, 19 (10), 3890-3905.

52. Kozak, R. P.; Royle, L.; Gardner, R. A.; Bondt, A.; Fernandes, D. L.; Wuhrer, M., Improved nonreductive O-glycan Release by Hydrazinolysis with Ethylenediaminetetraacetic Acid Addition. *Anal Biochem* 2014, 453, 29-37.
53. Song, X.; Ju, H.; Lasanajak, Y.; Kudelka, M. R.; Smith, D. F.; Cummings, R. D., Oxidative Release of Natural Glycans for Functional Glycomics. *Nat Methods* 2016, 13 (6), 528-534.
54. An, H. J.; Gip, P.; Kim, J.; Wu, S.; Park, K. W.; McVaugh, C. T.; Schaffer, D. V.; Bertozzi, C. R.; Lebrilla, C. B., Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells. *Mol Cell Proteomics* 2012, 11 (4), M111 010660.
55. Amicucci, M. J.; Nandita, E.; Galermo, A. G.; Castillo, J. J.; Chen, S.; Park, D.; Smilowitz, J. T.; German, J. B.; Mills, D. A.; Lebrilla, C. B., A Nonenzymatic Method for Cleaving Polysaccharides to Yield Oligosaccharides for Structural Analysis. *Nat Commun* 2020, 11 (1), 3963.
56. Bereman, M. S.; Williams, T. I.; Muddiman, D. C., Development of a nanoLC LTQ Orbitrap Mass-spectrometric Method for Profiling Glycans Derived from Plasma from Healthy, Benign Tumor Control, and Epithelial Ovarian Cancer Patients. *Analytical Chemistry* 2009, 81 (3), 1130-1136.
57. 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.
58. 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. *Analytical Chemistry* 2019, 91 (8), 5433-5445.
59. Tarasova, I. A.; Masselon, C. D.; Gorshkov, A. V.; Gorshkov, M. V., Predictive Chromatography of Peptides and Proteins as a Complementary Tool for Proteomics. *Analyst* 2016, 141 (16), 4816-4832.

Chapter II

N-Glycomic Analysis of the Cell Shows Specific Effects of Glycosyl Transferase Inhibitors

Abstract

Glycomic profiling methods were used to determine the effect of metabolic inhibitors on glycan production. These inhibitors are commonly used to alter the cell surface glycosylation. However, structural analysis of the released glycans has been limited. In this research, the cell membranes were enriched and the glycans were released to obtain the *N*-glycans of the glycocalyx. Glycomic analysis using liquid chromatography–mass spectrometry (LC–MS) equipped with a PGC chip column was used to profile the cell membrane glycocalyx. Glycans from untreated cells were compared to those from cells treated with inhibitors, including Kifunensine, which inhibits the formation of complex- and hybrid-type structures, 2,4,7,8,9-Penta-O-acetyl-*N*-acetyl-3-fluoro- β -D-neuraminic acid methyl ester for sialylated glycans, 2-deoxy-2-fluorofucose, and 6-alkynyl fucose for fucosylated glycans. Kifunensine was the most effective, converting nearly 95% of glycans to high-mannose types. The compound 6-alkynyl fucose inhibited some fucosylation but also incorporated into the glycan structure. Proteomic analysis of the enriched membrane for the four inhibitors showed only small changes in the proteome accompanied by large changes in the *N*-glycome for Caco-2. Future works may use these inhibitors to study the cellular behavior associated with the alteration of glycosylation in various biological systems, e.g., viral and bacterial infection, drug binding, and cell–cell interactions.

Introduction

The glycocalyx is the carbohydrate component of the cell membrane composed of glycans on glycoconjugates such as protein and lipids. *N*-Glycans are bound to asparagine on proteins and comprise the largest component of the glycocalyx¹. They are classified into three principal types: high-mannose, complex, and hybrid types. High-mannose structures are produced earlier in the

glycosylation process, while complex- and hybrid-type structures are produced later^{2,3}. Complex- and hybrid-type structures are further decorated by other monosaccharides, specifically, galactoses, fucoses, and sialic acids (*N*-acetyl-5-neuraminic acid or Neu5Ac). These additions play important roles in cell–cell interactions, including immune response, infection (viral and bacterial), and cell–cell-adhesion⁴⁻⁶. The variations in compositions, linkages (with anomeric character), and regiochemistry can have profound effects on protein function. For example, the absence of core fucose is generally lethal in humans⁷. Similarly, sialic acids are important structures for binding viruses and bacteria leading to infection^{8,9}.

Another important function of glycosylation is regulating protein turnover rates. It has been shown that IgGs decorated with sialic acid have longer half-lives than those without it, suggesting that glycosylation can affect protein turnover rates^{10,11}. We previously demonstrated the effect of glycosylation on protein turnover rates. In the study, we measured the turnover rate by feeding cells with isotope-labeled monosaccharide and measured the rate of incorporation of the exogenous sugars on the cell surface. It was shown that proteins glycosylated with high-mannose had faster turnover rates than proteins with complex-type structures¹².

Past studies have shown that the use of mutations to induce defects in glycosylation was difficult or caused unpredictable outcomes^{13,14}. The knockout of a specific glycosyltransferase can lead to another transferase taking over its role. Additionally, the removal of transferase genes, for example, FUT8, is lethal, making it more difficult to study in animals¹⁵. The development of metabolic glycosylation inhibitors helps overcome these limitations. Using metabolic glycosylation inhibitors is advantageous because small molecules are readily taken up into the cells. Additionally, these metabolic inhibitors make it possible to study animals since the amount needed to inhibit glycosylation is not lethal to the animals. These inhibitors enable the exploration

of fundamental questions regarding the decoration of specific monosaccharides on glycans. For example, fundamental questions regarding the presence of fucose or sialic acids have been examined using inhibitors. It has been shown that the downregulation of fucose on glycans using 2-fluoro-L-fucose (2FF), a fucosyltransferase inhibitor, suppresses the proliferation and migration of the HepG2 cell. The study showed that 2FF targeted fructosyltransferase 8, which is responsible for core fucosylation¹⁶. Additionally, 6-alkynyl fucose (6AF) is another known fucose inhibitor. It inhibits the protein FX, which is responsible for converting mannose to fucose¹⁷. It has been known that inhibition with 6AF has halted cell migration and invasion in hepatoma cells¹⁸. The reduction of sialic acid using 2,4,7,8,9-Penta-*O*-acetyl-*N*-acetyl-3-fluoro- β -D-neuraminic acid methyl ester (3FS) causes the impairment of adhesion, migration, and in vivo tumor growth on the B16F10 cell. The sialic acid inhibitor is suspected to bind to sialyltransferases after 3FS is converted to CMP-P-3FS¹⁹. Likewise, the increase in high-mannose glycans using a mannosidase I inhibitor, Kifuensine (Kif), has been shown to increase the metastatic characteristics of CCA cells²⁰. Though these inhibitors have been used in many studies, the alteration of the glycocalyx caused by the glycosylation inhibitor has not been well characterized using newer glycomics methods. Additionally, the effects on the expression of proteins in the glycocalyx by the presence of these inhibitors are little known.

The recent and general interest in the function of the glycocalyx has led to new tools to characterize them^{21,22}. Within the last decade, nano-Liquid Chromatography–Mass Spectrometry (nanoLC–MS)-based methods have been developed to characterize the glycocalyx. Using nanoflow liquid chromatography is advantageous because its high sensitivity allows for the detection of low abundant glycans. The coupling with mass spectrometry allows structural characterization using fragmentation patterns. One method we developed to study the glycome on

the cell surfaces uses a nanoflow liquid chromatography–quadrupole time-of-flight MS (nanoLC–QTOF MS). These methods have already been used to study glycans on human milk oligosaccharides as well as free glycans in bodily fluids²³.

In this research, we employed inhibitors to affect the cell surface glycosylation. We used N-glycans to monitor these changes and quantitate the change expressed by the inhibitors. The change caused by inhibition was monitored on three different cell lines (Caco-2, A549, and PNT2), and the changes in *N*-glycan profile were performed on a nanoLC-chip–QTOF MS. Additionally, using a nanoLC–Orbitrap MS, we performed proteomic analysis on Caco-2 to monitor the changes in protein expression that may be caused by altering the cell surface glycosylation.

Materials and Methods

Materials

Caco-2 cell lines, A549 cell lines, and Eagle’s Minimum Essential Medium (EMEM) were obtained from the American Type Culture Collection (ATCC), Manassas, VA, USA.

Dithiothreitol (DTT), iodoacetamide (IAA), and a human immortalized prostate PNT2 cell line were purchased from Sigma-Aldrich, St. Louis, MO, USA. Phosphate-Buffered Saline (PBS), Roswell Park Memorial Institute (RPMI) 1640 medium, fetal bovine serum (FBS), and penicillin were purchased from ThermoFisher Scientific, Waltham, MA, USA. Sequencing Grade Modified Trypsin was purchased from Promega. Glycosylation inhibitors Kifunensine, 2-deoxy-2-fluoro-L-fucose, 6-alkynyl fucose, and 2,4,7,8,9-Penta-*O*-acetyl-*N*-acetyl-3-fluoro- β -D-neuraminic acid methyl ester were purchased from Carbosynth, San Diego, CA, USA.

Cell Line Culture

Cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and cultured in their respected media supplemented with 10% fetal bovine serum and 1%

penicillin incubated at 37 °C in an atmosphere of 5% CO₂. Media were replaced every 48 h. Caco-2 cells were treated with 50, 100, 200, and 400 μM of 3FS, 2FF, 6AF, and 25, 50, 100, and 200 μM of Kif (as a positive control) once the cells reached 40% confluency. All other cells were treated with 200 μM of 3FS, 2FF, 6AF, and 100 μM of Kif. All compounds were dissolved in dimethyl sulfoxide (DMSO) with a final treatment concentration of 0.1% (v/v) for all inhibition experiments. Caco-2 was also treated with media containing 0.1% DMSO, which resulted in no difference in the profile between the 0.1% DMSO-treated cells to the control cell. After 72 h, the cells were washed with PBS and pelleted for N-glycan release or trypsin digestion and subsequent LC–MS analysis.

Cell Membrane Extraction

Extraction protocols were described previously and applied here with slight modifications²⁴. Cell pellets were reconstituted in a homogenization buffer containing 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.5), 0.25 M sucrose and protease inhibitors (Calbiochem/EMD Chemicals) at a 1:100 ratio. Using a probe sonicator (Qsonica, Newtown, CT, USA), cells were lysed with five alterations of on and off pulses in 5 and 10 s time intervals. Cellular debris and mitochondrial fractions were pelleted by centrifugation at 2000× g for 10 min. Supernatants were transferred to perform ultracentrifugation at 200,000× g for 45 min at 4 °C. Pelleted plasma membranes were then resuspended in 500 μL of 0.2 M Na₂CO₃ and 500 μL of nanopure water followed by two series of ultracentrifugation at 200,000× g for 45 min to wash off cytoplasmic and endoplasmic reticulum fractions.

N-Glycan Release

Enzymatic releases of *N*-glycans were adapted from a previously developed protocol where solid-phase extraction (SPE) was performed to enrich *N*-glycans²⁵. Samples were reconstituted in 100 μ L of 100 mM ammonium bicarbonate (ABC) and 5 mM DTT and heated at 100 °C for 1 min to denature proteins. Samples were cooled at room temperature followed by the addition of 2 μ L of 500,000 U/mL peptide *N*-glycosidase F (PNGase F), microwaved at 60 °C for 10 min to accelerate *N*-Glycans release. Samples were incubated for 18 h at 37 °C to hydrolyze the *N*-Glycans. The reaction was quenched with 350 μ L of water followed by ultracentrifugation at 200,000 \times *g* to separate the deglycosylated proteins and the *N*-glycans. *N*-glycans in the supernatant were collected, cleaned with porous graphited carbon (PGC) SPE, dried completely, and reconstituted in 30 μ L of water prior to analysis with Agilent 6200 series nano liquid chromatography chip–quadrupole time-of-flight mass spectrometry (nanoLC Chip–QTOF MS).

NanoLC Chip–QTOF MS Analysis

N-glycans analysis were described previously and applied here with slight modifications²⁶. *N*-glycans were separated using a nanoLC Chip–QTOF MS with a 65 min run time, where *N*-glycans begin to elute between 15 and 35 min. *N*-glycan separation was conducted with a binary solvent system, where solvent A was composed of 3% (*v/v*) acetonitrile (ACN) and 0.1% (*v/v*) formic acid (FA) in water, and solvent B was composed of 90% (*v/v*) ACN and 1% (*v/v*) FA in water. Samples were injected into an Agilent PGC microfluidic chip, which consisted of a 40 nL enrichment and a 75 μ m \times 43 nm analytical column, both with a particle size of 5 μ m. The gradient sequence for the run was: 0–2.5 min, 0% B; 2.5–20 min, 16% B; 20–35 min, 58% B; 35–40 min, 100% B; 40–50 min, 100% B; 50.01–65 min, 0% B with a flow rate of 0.3 μ L/min. Mass range of *m/z* 600–2000 with spectra were measured 1.5 s per

spectrum in positive ionization mode. Reference mass m/z 1221.991 were used to correct mass inaccuracies. Quantification of the cell surface glycosylation changes were determined using MassHunter software (Agilent, Santa Clara, CA, USA, version B.07.00). The software identifies molecular features and provides a chromatographic peak volume, which provides relative intensities. The intensities are used directly to provide relative abundances. *N*-Glycan tandem data were analyzed with GlycoNote (<https://github.com/MingqiLiu/GlycoNote>, Last accessed on 12 May 2021). GlycoNote structures were then validated manually. K-mean cluster analysis using JMP software (SAS, Cary, NC, USA, version Pro 15) was also applied to *N*-glycans results.

Leica TCS SP8 STED 3X Fluorescence Imaging

PNT2 cells were cultured in a 35 mm glass bottom dish (ibidi) until 20% confluency. Cells were treated with 200 μ M of 6AF for 72 h at 37 °C in an atmosphere of 5% CO₂. After 3 days, cells were fixed with 1% paraformaldehyde, “click” with a fluorescent tag 7-Azido-4-Methylcoumarin from Sigma-Aldrich with excitation at 387 nm and emission at 470 nm, and membrane stained with Cellmark dye (ThermoFisher, Waltham, MA, USA) with excitation at 522 nm and emission at 535 nm. Images were taken at VetMed Advance Imaging Facility under 100 \times magnification with oil submersion.

Trypsin Digestion

Enzymatic digestion of membrane proteins was adapted from previous procedures followed by peptide desalting using 500 mg of C18 SPE cartridges²⁷. The cell pellet was sonicated with 60 μ L of 8M urea for 15 min followed by the addition of 2 μ L of 550 mM DTT incubated at 55 °C for 50 min. After incubation, 4 μ L of 450 mM of iodoacetamide (IAA) was added and placed in the dark for 20 min. An aliquot of 420 μ L of 50 mM ABC was added to each sample with the

addition of 2 µg of trypsin that was reconstituted in 50 mM ABC. Samples were then incubated at 37 °C for 18 h. For proteomic analysis, peptides were desalted using C18 SPE cartridges. The peptide concentration for proteomic analysis was determined with bicinchoninic acid assay (BCA) peptide assay (ThermoFisher, Waltham, MA, USA) prior to injection in the Orbitrap Fusion Lumos nanoLC–MS/MS instrument.

Proteomic Data Analysis

The Human FASTA database was acquired from UniProt. The Raw files and FASTA file were then inputted into Byonic software version v3.11.3 for proteomic analysis followed by extract ion chromatogram (EIC) using Byologic software (Protein Metrics, Cupertino, Ca, USA, version v3.11.3). Posterior error probability (PEP) smaller than 0.01, Score > 100, peptide length > 5, and 1% false discovery rate (FDR) were applied for proteomic data analysis. Multiple t tests were performed using GraphPad Prism software (GraphPad, San Diego, CA, USA version 8).

Results

The workflow for the glycomic characterization of the cell lines is shown schematically in **Figure 2.1**. It involves the harvesting and lysing of the cells followed by a series of ultracentrifugation steps to enrich the cell membrane fractions. The enriched fractions are subjected to *N*-glycan release using the enzyme PNGase F. NanoLC–MS using a PGC stationary phase and QTOF mass analyzer produces a glycan profile with isomer separation. In the *N*-glycomic profile, we generally identify over 300 structures. The membrane enrichment has been validated in previous publications²⁸.

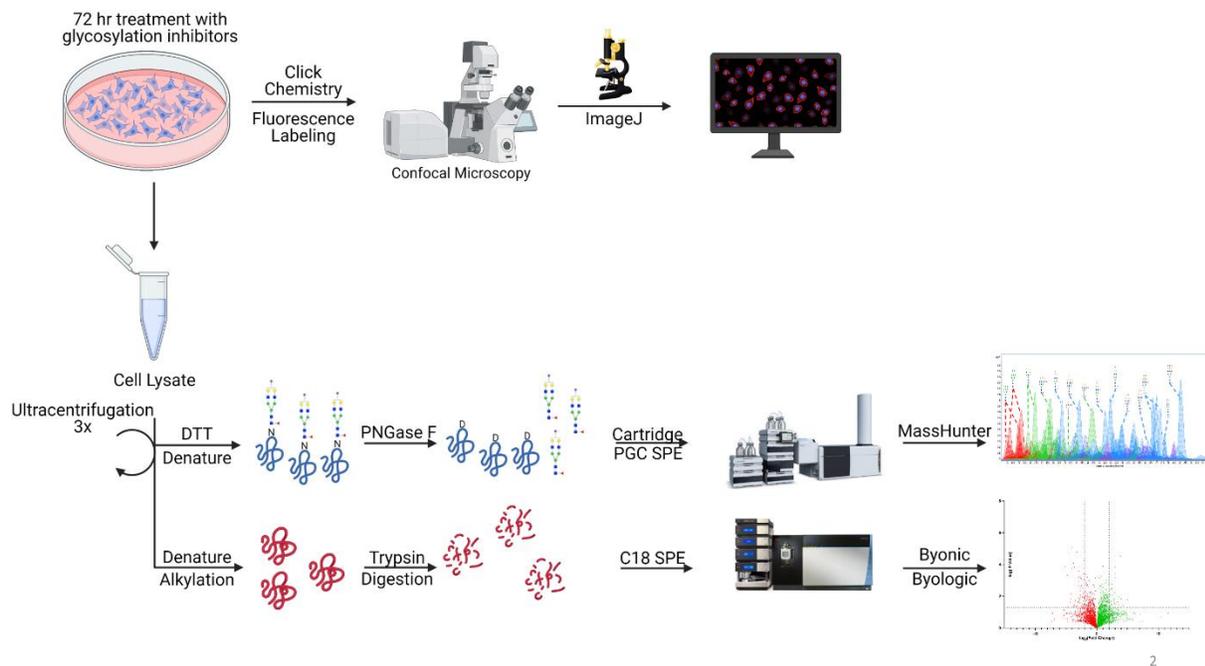


Figure 2.1 The schematic workflow for characterization of N-glycomics, proteomics, and confocal microscopy. Glycosylation inhibitors were added to the supernatant and treated for 72 h followed by cell membrane extraction and analysis. Click chemistry was performed after 72 h of treatment followed by imaging to determine 6-alkynyl fucose (6AF) incorporation. Schematic workflow was created with BioRender.com last accessed 17 August 2021.

N-Glycan Profile of Cell Membranes

The *N*-glycome profile for Caco-2, a colon epithelial carcinoma line, was determined and yielded over 200 glycans (including isomers) (**Figure 2.2A**). The untreated cells yielded glycans composed of high-mannose- (18.0%), complex- (47.3%), and hybrid-type structures (34.7%). With regard to the complex glycans, they were composed of mono- (0.3%), bi- (2.0%), tri- (6.5%), and tetra-antennary (29.5%) structures. Other structural features included bisecting *N*-acetylglucosamine (GlcNAc) (9.0%). The complex and hybrid compounds were further separated into sialylated (3.8%), fucosylated (33.7%), and sialofucosylated (42.7%). The

sialylated glycans were mainly singly (42.6%), doubly (12.9%), triply (3.5%) and quadruply sialylated (0.4%) glycans. Similarly, the fucosylated glycans were composed of singly (36.7%), doubly (24.8%), triply (13.2%), quadruply (6.9%) and quintuply fucosylated (1.1%) glycans. These results were consistent with previously published values for the undifferentiated Caco-2 cell line, which similarly yielded significantly higher amounts of complex-type structures with high levels sialofucosylation. There were over 300 structures composed of nearly 100 compositions with only 16 representing nearly 50% of the intensities (**Supplementary Figure 2.1**). The most abundant composition was Hex₅HexNAc₅Fuc₁NeuAc₁.

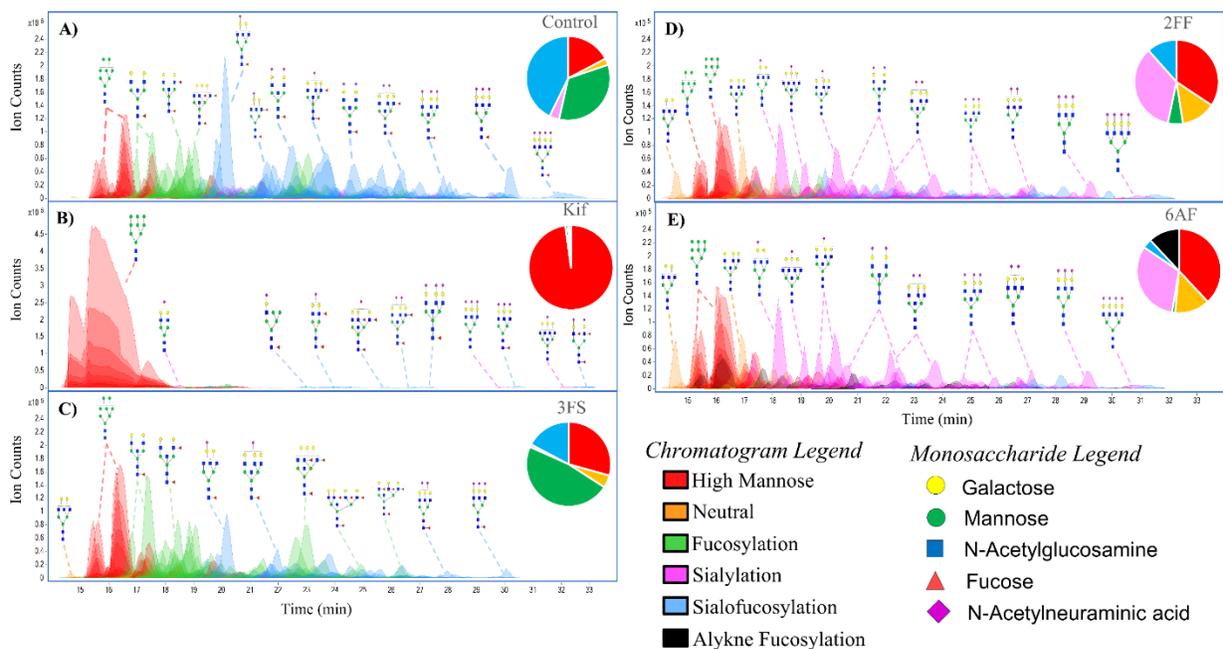


Figure 2.2 Extracted compound chromatograms of Caco-2, showing (A) control, (B) Kifunensine (Kif), (C) 2,4,7,8,9-Penta-O-acetyl-N-acetyl-3-fluoro-b-D-neuraminic acid methyl ester (3FS), (D) 2-deoxy-2-fluorofucose (2FF), and (E) 6-alkynyl fucose. Peaks are colored by glycan subtypes and annotated with the schematic representation of the glycan structures. Monosaccharide

notations follow the Symbol Nomenclature for Glycans (SNFG) system²⁹ and N-glycans were drawn with GlycoWorkbench version 2.1³⁰.

The inhibitors were introduced to the media with 0.1% DMSO. To ensure that DMSO had little effect on the cell surface glycan expression, a comparison of DMSO-treated cells with untreated cells was performed. The difference was barely observed (**Supplementary Figure 2.2A**), which suggested that DMSO had no effect and did not alter the membrane glycome. To determine the reproducibility, triplicate experiments were performed. **Supplementary Figure 2.2B** shows the *N*-glycan profiles of Caco-2 cells treated with 2FF inhibitor performed in triplicates. The coefficient of variation (CV) for compounds varied less than 20%. The variations in the *N*-glycan subtypes averaged were below 20%.

The effect of the inhibitor concentration was monitored. Concentrations between 25 and 400 μM were introduced to the supernatant and allowed to incubate for several hours. The glycan profiles were obtained for each experiment. For Kif, 100 μM concentrations and 72 h incubation proved optimal. For the other inhibitors, higher concentrations were needed at 200 μM and 72 h (**Supplementary Figure 2.3**).

The addition of Kif produced the largest effect on the glycocalyx. The LC-MS chromatogram yielded primarily high-mannose structures (over 95% in total abundances) (**Figure 2.2B**). The remaining compounds in abundances are residual complex-type *N*-glycans. In comparison, the untreated cells contained only 18.0% high-mannose structures.

The compound 3FS is a known sialyltransferase inhibitor. The sialic acid analog is incorporated into the salvage pathway and converted to the active form of the sugar by attaching cytidine-5'-monophosphate (CMP), resulting in the nucleotide sugar CMP-3FS^{31,32}. The active

form of the sugar then inhibits the transferase, depleting the cells of sialylated glycans. **Figure 2.2C** shows the *N*-glycan profile for Caco-2 treated with 3FS. Firstly, there was no incorporation of 3FS into the cell surface *N*-glycome. The 3FS-treated Caco-2 cells were composed of high-mannose (29.9%), complex (37.5%), and hybrid-type (32.6%) structures. Sialylated and sialofucosylated glycans decreased by 28% while fucosylated glycans increased by 14% compared to the untreated cell. The increase in fucosylated glycans is due to the loss of sialylated decoration in the previously sialofucosylated species. For example, the compound Hex₅HexNAc₅Fuc₁ increased from 3.3% to 6.0% upon treatment with 3FS. The increase is caused by the loss of sialic acids from the three sialylated analogs, namely, Hex₅HexNAc₅Fuc₁NeuAc₁, Hex₅HexNAc₅Fuc₁NeuAc₂, and Hex₅HexNAc₅Fuc₁NeuAc₃. However, it should be noted that the total loss of sialylated glycans was not observed. Among the fucosylated species, Hex₅HexNAc₅Fuc₁NeuAc₁ was still the most abundant sialofucosylated glycan.

The fucose transferase inhibitor 2FF is converted with guanosine 5'-diphosphate (GDP) to GDP-2FF through the fucose salvage pathway³³. This active sugar inhibits fucosyltransferases and led to the profile shown in **Figure 2.2D**. We examined the glycan profile and determined that 2FF was similarly not incorporated into the *N*-glycans. The *N*-glycome was composed of 34.7% high-mannose, 39.7% complex, and 25.6% hybrid. Fucosylated glycans decreased from 76% to 17% after 2FF treatment. An increase in the sialylated glycans was observed from 3.8% to 35.1%. For example, the structure Hex₅HexNAc₅NeuAc₁ increased from 0.5% to 6.5% likely due to the decrease in fucosylated species, including Hex₅HexNAc₅Fuc₁NeuAc₁, Hex₅HexNAc₅Fuc₂NeuAc₁, and Hex₅HexNAc₅Fuc₃NeuAc₁. An important observation in comparing the two inhibitors, 2FF and 3FS, was that they appeared to behave independently. It

seems that the decoration of fucose and the decoration of sialic acid on sialofucosylated glycans are not dependent on one another.

The compound 6AF was known to inhibit the de novo pathway through interactions with 3,5-epimerase-4-reductase, thereby diminishing fucose incorporation. This pathway converts mannose into fucose, forming the activated reagent GDP-fucose¹⁸. Unlike the other inhibitors, examination of the N-glycome profile showed that 6AF was incorporated into specific glycans (**Figure 2.2E**). The N-glycome profile yielded 38.1% high-mannose, 37.7% complex, and 24.2% hybrid. Complex and hybrid glycans were composed of 1.2% fucosylated, 31.8% sialylated, and 3.6% sialofucosylated. An additional 11.9% was obtained for the 6AF incorporated species corresponding to 6AF-fucosylated and 6AF-sialofucosylated glycans combined. In general, 6AF behaves similarly to 2FF in that total fucosylation was decreased. However, incorporation of the inhibitor was also observed for 6AF.

The incorporation of 6AF was further confirmed by MS/MS analysis (**Supplementary Figure 2.4**). However, no attempt was made to differentiate between core and antennary fucosylation. It is noted that only one 6AF was incorporated at most, with the most abundant being Hex₆HexNAc₇6AF₁NeuAc₃. For PNT2, the inhibition created a nearly 40% drop in fucosylation, with the remaining fucosylated species containing between 20 and 30% 6AF and the rest with native fucose. To validate the incorporation, we probed cell lines with Click chemistry using a chromophoric tag. A control and 6AF-treated PNT2 cells were reacted for one hour with 7-Azido-4-Methylcoumarin followed by staining with CellMask™ Deep Red (cell plasma membrane). The merged images of 6AF and CellMask™ confirmed the incorporation of 6AF both on the cell plasma membrane and Golgi (**Figure 2.3**). The incorporation of 6AF under

similar conditions was reported previously¹⁷. The inhibitor should therefore be used with caution in assessing fucose function.

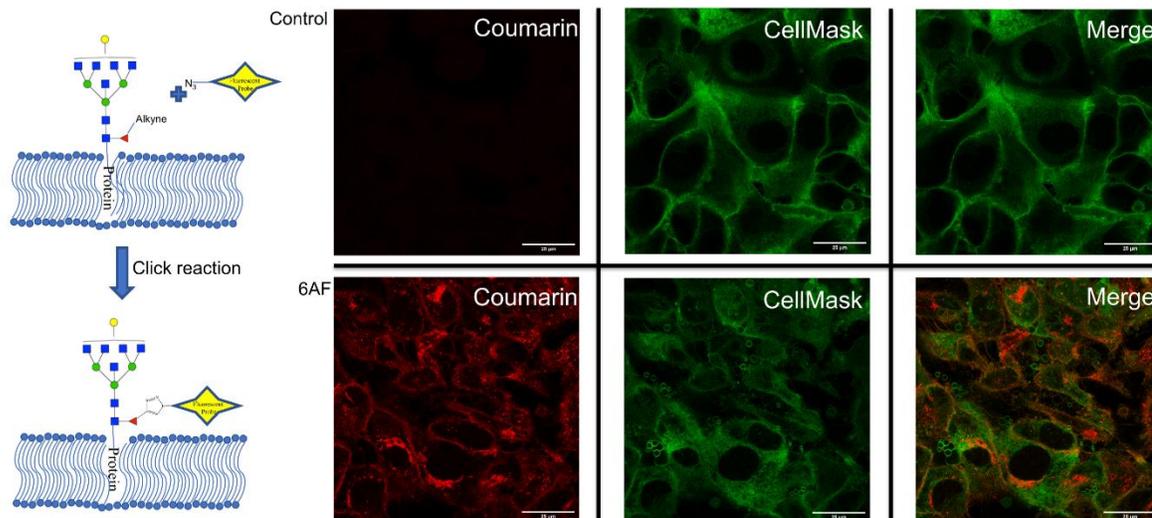


Figure 2.3 Confocal microscopy of PNT2 cells were the top cells without 6AF treatment (control) and the bottom cells with 6AF treatment. Cells were stained with CellMask™ (green) and 7-Azido-4-Methylcoumarin (red) (ThermoFisher, Waltham, MA, USA). The merge section is the overlap between CellMask™ and 7-Azido-4-Methylcoumarin.

To observe the general behavior of the inhibitors, each one was similarly examined with additional cell lines, namely, A549 (lung epithelial carcinoma) and PNT2 (normal prostate epithelial). These cell types were chosen to study the effect of each inhibitor on different cell types. The comparison of the results for Caco-2, A549, and PNT2 (**Figure 2.4**) was presented according to their major glycan types. A549 and PNT2 behaved similarly when treated with Kif yielding high-mannose glycans with abundances of 96% and 85%, respectively. When treated with 3FS, the cells lost most of their sialylated glycans, with A549 having only 8% sialylated glycans after treatment compared to 50% before, which is very similar to PNT2 with 6%

compared to 47%. Fucosylated glycans in PNT2 cells treated with 2FF decreased by 7%, while A549 decreased by 24%. PNT2, when treated with 6AF, lost more fucosylated glycans compared to 2FF with 33% loss and a 9% incorporation of 6AF. No incorporation of 2FF was observed in any of the cell lines. A549 lost 17% of its fucosylated glycans with a 7% incorporation of 6AF.

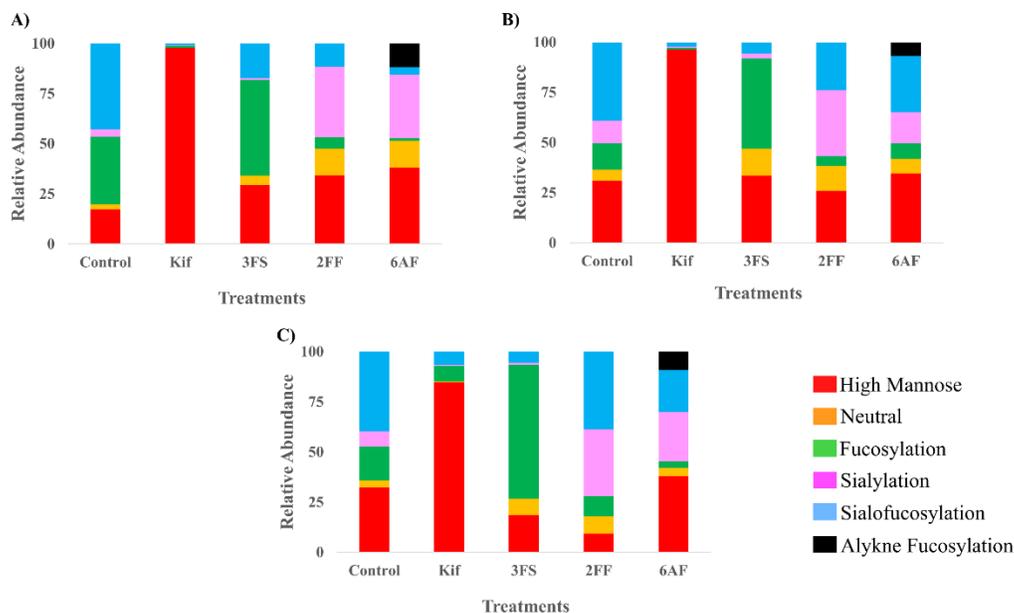


Figure 2.4. The relative abundance profile of each glycosylation inhibitor with its control counterpart plotted on stack columns with (A) Caco-2, (B) A549, and (C) PNT2 cell lines. Stack column plots are color coded by glycan subtypes.

K-mean clustering analysis was performed on the inhibition dataset with all three cell lines. K-mean clustering is a type of unsupervised machine learning where random centroids are generated. The algorithm then calculates the Euclidian distance between a point and the nearest centroid. Afterward, all the points nearest to the centroid are averaged and the centroid position is shifted to the average distance. This is done for all centroids, and it iterates until convergence. Through this process, each treatment was clustered together with one another regardless of which

cell lines were being treated (**Supplementary Figure 2.5**). The orange cluster contains cells treated with Kif. This cluster had the tightest pack, suggesting that the inhibition effects were most similar to each other regardless of the cell type, while 6AF-treated cells (light green cluster) had the most variation between the cells based on the spacing in the clustering. Though each cell line had slight variation in the results, the overall effect of the inhibitors was independent of the cell itself.

Cell Surface Proteomic with Inhibitors

To determine whether the glycan modification by the inhibitors altered protein expression, proteomic analysis was performed on enriched cell membrane. A volcano plot (**Figure 2.5**) was constructed to determine the similarities between the proteins in each treatment group compared to the untreated control group.

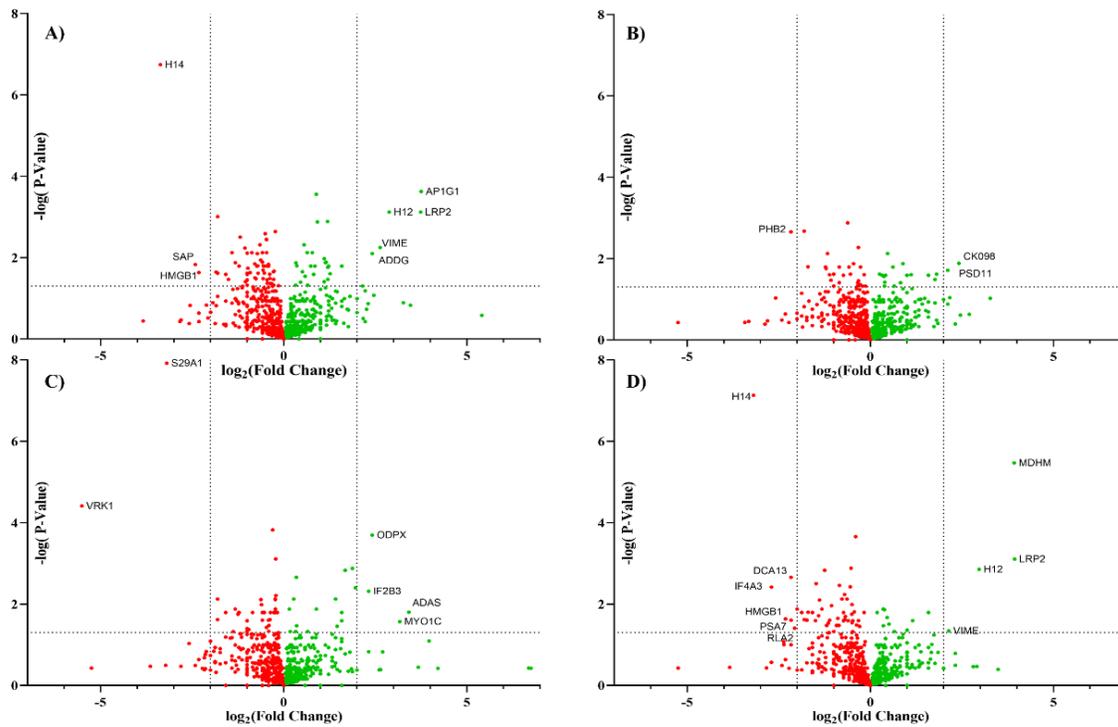


Figure 2.5 The volcano plot represents the proteomic data from Caco-2 comparing control to an inhibition treatment including (A) Kif, (B) 3FS, (C) 2FF, and (D) 6AF. Each data point represents a protein with the downregulation (red) and upregulation (green).

Biological triplicates of Caco-2 cells treated with each inhibitor were compared to untreated triplicate cells using raw XIC values extracted from Byologic software. The resulting data were sorted by peptide length less than 5 and score less than 100 to remove potential missed cleavages sequence. Each protein was represented by the most abundant peptide signal. The thresholds for significant proteins were p-values less than or equal to 0.05 and \log_2 fold change less than -2 or greater than 2.

Among all the treatments, the largest change in the glycomic profile was observed by the addition of kifunensine. Kif converted the glycans from complex-type structures to high-mannose structures with the major abundance corresponding to Man₉. Proteomic analysis of both Kif-treated cells and the control yielded around 1100 proteins for each treatment. There were over 1000 proteins that were conserved in both groups, and 9 proteins were identified as significant using the above criteria. Among those nine proteins, only five were cell surface proteins. The proteomic results show little variations in the proteins identified between the control and Kif-treatment. Based on these results, there is apparently little change in protein expression even with large dramatic changes in glycan expression, suggesting that glycosylation does not appear to affect the membrane protein composition significantly.

The comparison of the other treatments yielded similar results with minor variations between the treatments. They all yielded around 1100 proteins, with approximately 1000 proteins conserved between the treated cells and the untreated control cells. Inhibiting sialylation with 3FS or inhibiting with fucosylation yielded a smaller change in the membrane protein expression

compared to Kif. On the other hand, 6AF inhibition yielded more significant numbers of proteins than Kif-treated cells using the above criteria. There were 17 proteins that were determined to be significant in 6AF-treated cells. Even so, only a handful of proteins were determined to be cell surface proteins (LRP2, VIME, and HMGB1). This further suggests that the inhibition of cell surface glycosylation had little effect on protein expression.

Discussion

Glycomic profiling provides an extensive characterization of metabolic inhibitors and their efficacy. The efficacy of each inhibitor varied. Among them, Kif had the highest efficacy and was capable of removing all other glycan types besides high-mannose type. Kif is known to inhibit mannosidase I (Man 1A1), the enzyme necessary to trim high-mannose structures and produce complex- and hybrid-type structures³⁴. Thus, the action of this compound primarily yields the Man₉ glycan (Hex₉HexNAc₂). The conversion was high, up to 97% high-mannose. There were remaining complex-type structures in, for example, Caco-2, resulting in some 3% remaining sialofucosylated glycans. In general, none of the inhibitors yielded complete conversion, at least as detected by the method. There could be several factors that can leave residual glycans, with the most likely cause to be differences in protein turnover associated with the glycoform. We have previously shown that proteins on the cell membrane express differential turnover rates¹², with sialylated proteins having the lowest rates. Indeed, the remaining were generally complex-type glycans with sialylation. In general, sialylation is more difficult to suppress due the slower turnover and the greater number of sialyltransferase.

Cells treated with the sialic acid inhibitor 3FS decreased sialylation by approximately 60%. The sialic acid analog is incorporated into the salvage pathway and converted to the active form of the sugar by attaching CMP resulting in the nucleotide sugar CMP-3FS^{31,32}. The active form of

the sugar then inhibits the transferase, depleting the cells of sialylated glycans. Cells treated with fucose inhibitors 2FF and 6AF decreased fucosylated species by nearly 80%. 6AF had the additional complication of incorporating into the glycan. The efficacy of each inhibitor was consistent regardless of cell type shown in the N-glycome profile of different cell types and K-mean clustering analysis. Additionally, Caco-2 proteomic results indicated that each inhibitor had little effect on the proteins expressed on the cell surface membrane for this cell line. The results strongly suggest that altered glycosylation does not affect protein expression in the glycocalyx of Caco-2. Kif had the highest efficacy, but less than one percent of the cell surface proteomic showed a significant change in protein expression.

Conclusion

In summary, we were able to profile the effect of these glycosylation inhibitors on the glycocalyx using a nanoLC-chip-QTOF MS, overcoming the limitation of other non-structurally specific methods. Through this method of analysis, we were able to obtain a more qualitative analysis of each inhibitor's effect to give a more in-depth understanding of how each inhibitor alters glycosylation. Future studies can use this application to study the functional role of different decorations of glycosylation in other biological systems.

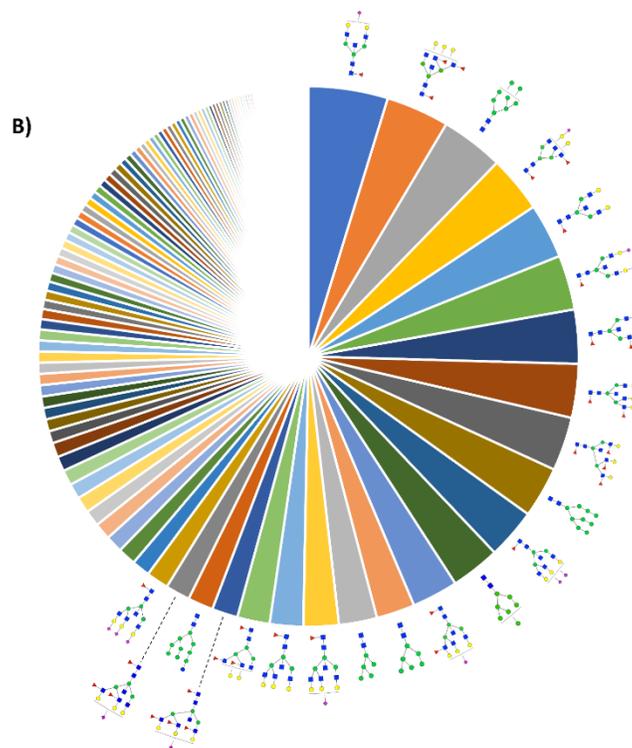
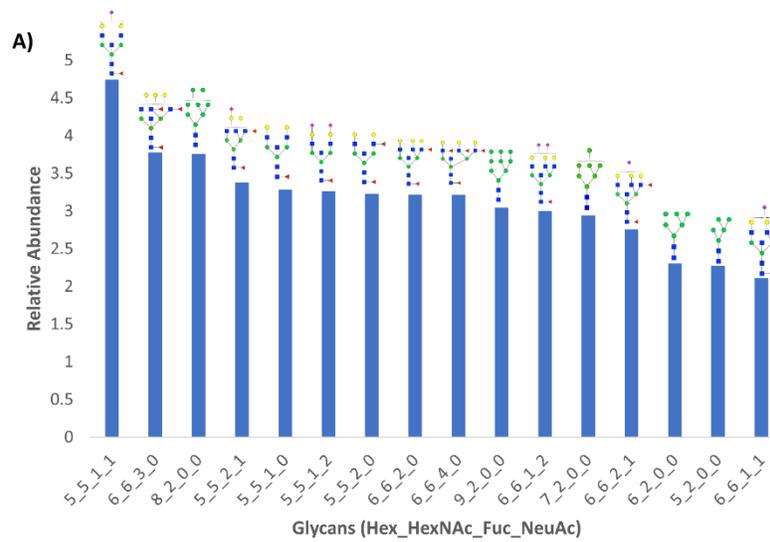
References

1. Schjoldager, K.T.; Narimatsu, Y.; Joshi, H.J.; Clausen, H. Global View of Human Protein Glycosylation Pathways and Functions. *Nat Rev Mol Cell Biol* 2020, 21, 729-749, doi:10.1038/s41580-020-00294-x.
2. Vervecken, W.; Kaigorodov, V.; Callewaert, N.; Geysens, S.; De Vusser, K.; Contreras, R. In Vivo Synthesis of Mammalian-Like, Hybrid-Type N-Glycans in *Pichia Pastoris*. *Appl Environ Microbiol* 2004, 70, 2639-2646, doi:10.1128/aem.70.5.2639-2646.2004.
3. Bieberich, E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. *Adv Neurobiol* 2014, 9, 47-70, doi:10.1007/978-1-4939-1154-7_3.
4. Lin, B.; Qing, X.; Liao, J.; Zhuo, K. Role of Protein Glycosylation in Host-Pathogen Interaction. *Cells* 2020, 9(4), 1022 doi:10.3390/cells9041022.
5. Marth, J.D.; Grewal, P.K. Mammalian Glycosylation in Immunity. *Nat Rev Immunol* 2008, 8, 874-887, doi:10.1038/nri2417.
6. Ohtsubo, K.; Marth, J.D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* 2006, 126, 855-867, doi:10.1016/j.cell.2006.08.019.
7. Haltiwanger, R.S. Fucose is on the TRAIL of Colon Cancer. *Gastroenterology* 2009, 137, 36-39, doi:10.1053/j.gastro.2009.05.010.
8. Baker, A.N.; Richards, S.J.; Guy, C.S.; Congdon, T.R.; Hasan, M.; Zwetsloot, A.J.; Gallo, A.; Lewandowski, J.R.; Stansfeld, P.J.; Straube, A.; et al. The SARS-COV-2 Spike Protein Binds Sialic Acids and Enables Rapid Detection in a Lateral Flow Point of Care Diagnostic Device. *ACS Cent Sci* 2020, 6, 2046-2052, doi:10.1021/acscentsci.0c00855.
9. Severi, E.; Hood, D.W.; Thomas, G.H. Sialic Acid Utilization by Bacterial Pathogens. *Microbiology (Reading)* 2007, 153, 2817-2822, doi:10.1099/mic.0.2007/009480-0.
10. Russell, A.; Adua, E.; Ugrina, I.; Laws, S.; Wang, W. Unravelling Immunoglobulin G Fc N-Glycosylation: A Dynamic Marker Potentiating Predictive, Preventive and Personalized Medicine. *Int J Mol Sci* 2018, 19, doi:10.3390/ijms19020390.
11. Shental-Bechor, D.; Levy, Y. Effect of Glycosylation on Protein Folding: A Close Look at Thermodynamic Stabilization. *Proc Natl Acad Sci U S A* 2008, 105, 8256-8261.
12. 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, 859-871, doi:10.1093/glycob/cwaa038.
13. Stanley, P. What Have We Learned from Glycosyltransferase Knockouts in Mice? *J Mol Biol* 2016, 428, 3166-3182, doi:10.1016/j.jmb.2016.03.025.
14. Freeze, H.H. Genetic Defects in the Human Glycome. *Nat Rev Genet* 2006, 7, 537-551, doi:10.1038/nrg1894.
15. Ng, B.G.; Xu, G.; Chandy, N.; Steyermark, J.; Shinde, D.N.; Radtke, K.; Raymond, K.; Lebrilla, C.B.; AlAsmari, A.; Suchy, S.F.; et al. Biallelic Mutations in FUT8 Cause a Congenital Disorder of Glycosylation with Defective Fucosylation. *Am J Hum Genet* 2018, 102, 188-195, doi:10.1016/j.ajhg.2017.12.009.
16. 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, 11563, doi:10.1038/s41598-017-11911-9.
17. Ma, C.; Takeuchi, H.; Hao, H.; Yonekawa, C.; Nakajima, K.; Nagae, M.; Okajima, T.; Haltiwanger, R.S.; Kizuka, Y. Differential Labeling of Glycoproteins with Alkynyl Fucose Analogs. *International Journal of Molecular Sciences* 2020, 21, 6007.
18. Kizuka, Y.; Nakano, M.; Yamaguchi, Y.; Nakajima, K.; Oka, R.; Sato, K.; Ren, C.T.; Hsu, T.L.; Wong, C.H.; Taniguchi, N. An Alkynyl-Fucose Halts Hepatoma Cell Migration and

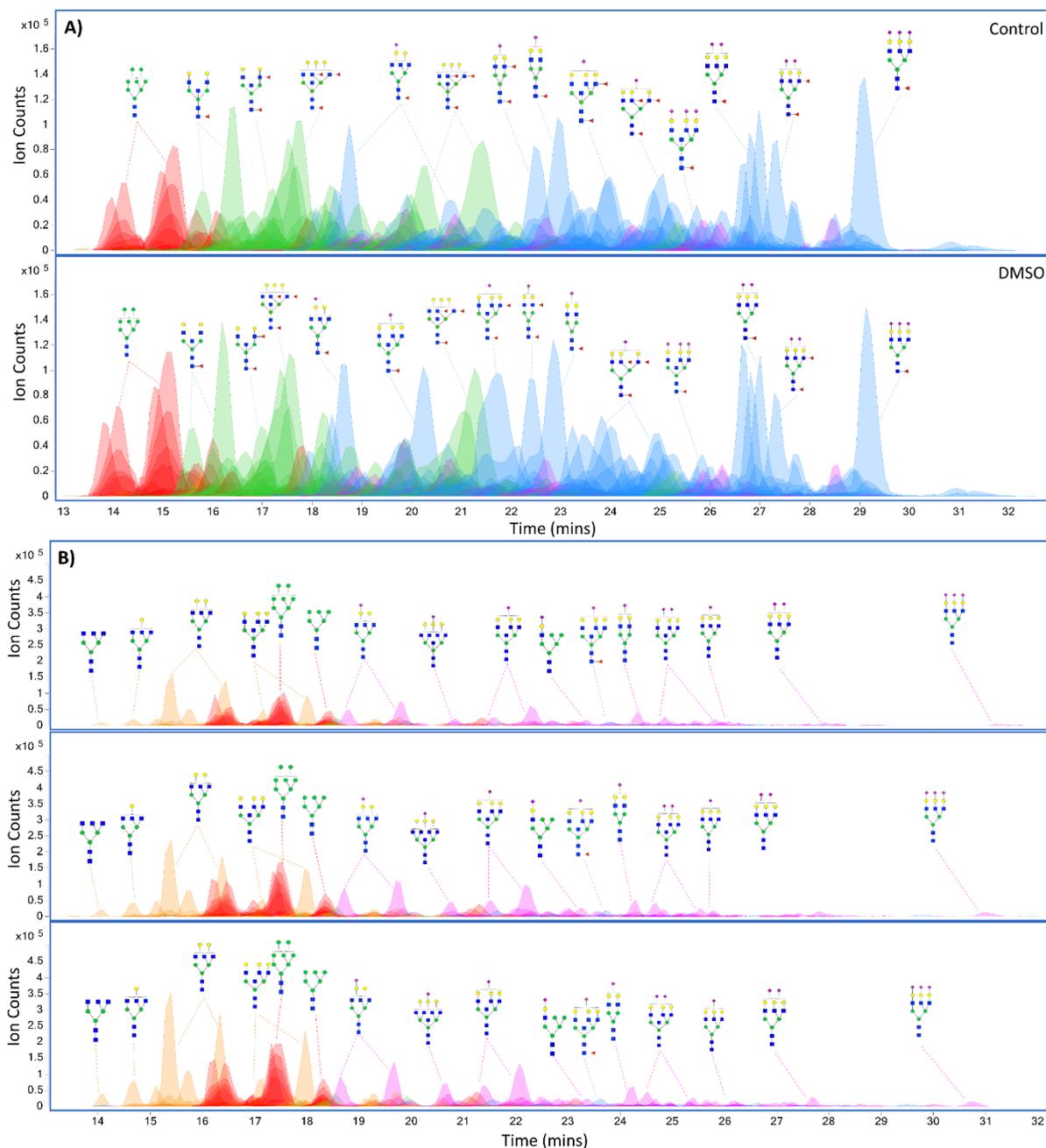
- Invasion by Inhibiting GDP-Fucose-Synthesizing Enzyme FX, TSTA3. *Cell Chem Biol* 2017, 24, 1467-1478 e1465, doi:10.1016/j.chembiol.2017.08.023.
19. Bull, C.; Boltje, T.J.; Wassink, M.; de Graaf, A.M.; van Delft, F.L.; den Brok, M.H.; Adema, G.J. Targeting Aberrant Sialylation in Cancer Cells Using a Fluorinated Sialic Acid Analog Impairs Adhesion, Migration, and In Vivo Tumor Growth. *Mol Cancer Ther* 2013, 12, 1935-1946, doi:10.1158/1535-7163.MCT-13-0279.
 20. Park, D.D.; Phoomak, C.; Xu, G.; Olney, L.P.; Tran, K.A.; Park, S.S.; Haigh, N.E.; Luxardi, G.; Lert-Itthiporn, W.; Shimoda, M.; et al. Metastasis of Cholangiocarcinoma is Promoted by Extended High-mannose Glycans. *Proc Natl Acad Sci U S A* 2020, 117, 7633-7644, doi:10.1073/pnas.1916498117.
 21. Kailemia, M.J.; Xu, G.; Wong, M.; Li, Q.; Goonatileke, E.; Leon, F.; Lebrilla, C.B. Recent Advances in the Mass Spectrometry Methods for Glycomics and Cancer. *Anal Chem* 2018, 90, 208-224, doi:10.1021/acs.analchem.7b04202.
 22. Ruhaak, L.R.; Xu, G.; Li, Q.; Goonatileke, E.; Lebrilla, C.B. Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem Rev* 2018, 118, 7886-7930, doi:10.1021/acs.chemrev.7b00732.
 23. Wu, S.; Tao, N.; German, J.B.; Grimm, R.; Lebrilla, C.B. Development of an Annotated Library of Neutral Human Milk Oligosaccharides. *J Proteome Res* 2010, 9, 4138-4151, doi:10.1021/pr100362f.
 24. Park, D.; Brune, K.A.; Mitra, A.; Marusina, A.I.; Maverakis, E.; Lebrilla, C.B. Characteristic Changes in Cell Surface Glycosylation Accompany Intestinal Epithelial Cell (IEC) Differentiation: High Mannose Structures Dominate the Cell Surface Glycome of Undifferentiated Enterocytes. *Mol Cell Proteomics* 2015, 14, 2910-2921, doi:10.1074/mcp.M115.053983.
 25. 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, 6271-6285, doi:10.1039/c8sc01875h.
 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, 2668-2704, doi:10.1038/s41596-020-0350-4.
 27. Li, Q.; Xie, Y.; Xu, G.; Lebrilla, C.B. Identification of Potential Sialic Acid Binding Proteins on Cell Membranes by Proximity Chemical Labeling. *Chem Sci* 2019, 10, 6199-6209, doi:10.1039/c9sc01360a.
 28. An, H.J.; Gip, P.; Kim, J.; Wu, S.; Park, K.W.; McVaugh, C.T.; Schaffer, D.V.; Bertozzi, C.R.; Lebrilla, C.B. Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of High Mannose Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells. *Mol Cell Proteomics* 2012, 11, M111 010660, doi:10.1074/mcp.M111.010660.
 29. Varki, A.; Cummings, R.D.; Aebi, M.; Packer, N.H.; Seeberger, P.H.; Esko, J.D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* 2015, 25, 1323-1324, doi:10.1093/glycob/cwv091.
 30. Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S.M. GlycoWorkbench: a Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. *J Proteome Res* 2008, 7, 1650-1659, doi:10.1021/pr7008252.
 31. Rillahan, C.D.; Antonopoulos, A.; Lefort, C.T.; Sonon, R.; Azadi, P.; Ley, K.; Dell, A.; Haslam, S.M.; Paulson, J.C. Global Metabolic Inhibitors of Sialyl- and Fucosyltransferases Remodel the Glycome. *Nat Chem Biol* 2012, 8, 661-668, doi:10.1038/nchembio.999.
 32. Suzuki, K.; Daikoku, S.; Son, S.H.; Ito, Y.; Kanie, O. Synthetic Study of 3-Fluorinated Sialic Acid Derivatives. *Carbohydr Res* 2015, 406, 1-9, doi:10.1016/j.carres.2014.12.010.

33. Tu, Z.; Lin, Y.N.; Lin, C.H. Development of Fucosyltransferase and Fucosidase Inhibitors. *Chem Soc Rev* 2013, 42, 4459-4475, doi:10.1039/c3cs60056d.
34. Elbein, A.D.; Tropea, J.E.; Mitchell, M.; Kaushal, G.P. Kifunensine, a Potent Inhibitor of the Glycoprotein Processing Mannosidase I. *Journal of Biological Chemistry* 1990, 265, 15599-15605, doi:10.1016/s0021-9258(18)55439-9.

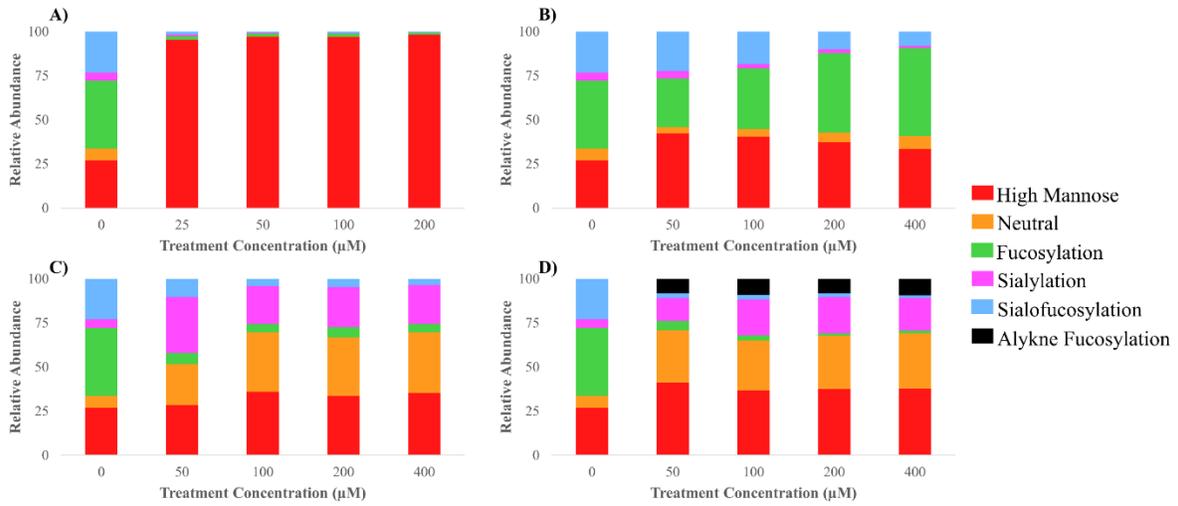
Supplementary Figures



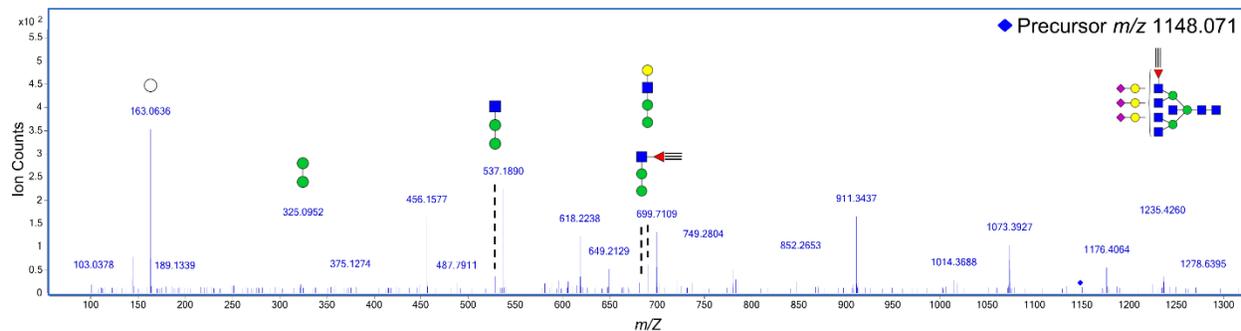
Supplementary Figure 2.1 Relative abundant N-glycans found in Caco-2 (A) the most abundant glycans and (B) are all the N-glycan identified with MassHunter software. N-Glycans were drawn with GlycoWorkbench.



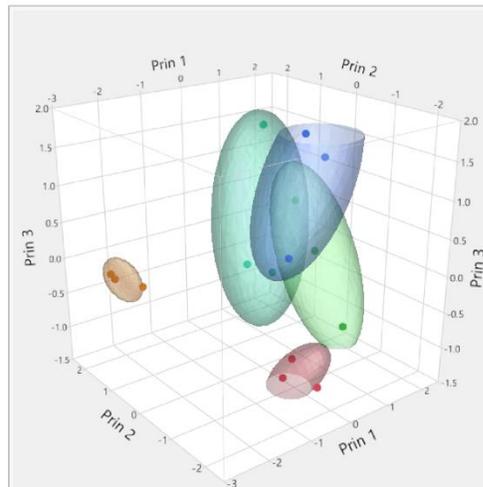
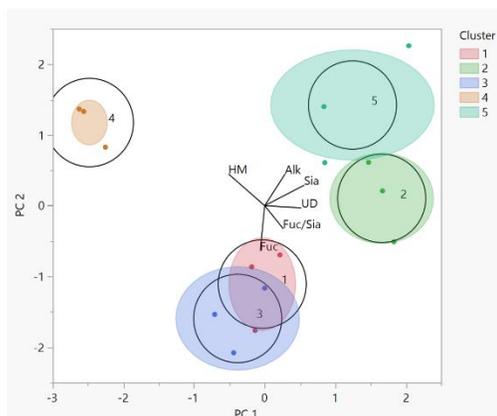
Supplementary Figure 2.2 Extracted compound chromatograms of Caco-2, showing (A) comparing untreated control to DMSO treated cells, (B) biological replicates of Caco-2 treated with 2-deoxy-2-fluorofucose. Peaks are colored by glycan subtypes and annotated with the schematic representation of the glycan structures. N-Glycans were drawn with GlycoWorkbench.



Supplementary Figure 2.3 The relative abundance of Caco-2 with different concentration of glycosylation inhibitors, showing (A) Kifunensine, (B) 3Fax-Neu5Ac, (C) 2-deoxy-2-fluorofucose, and (D) 6-alkynyl fucose. Concentration ranges from 25µM to 400µM and in each stack column plot are color coded by glycan subtypes.



Supplementary Figure 2.4 N-Glycomic MS/MS spectra of PNT2 treated with 6AF. This exhibits 6AF incorporation in Hex₆HexNAc₇6AF₁NeuAc₃. Monosaccharide notations follow the Symbol Nomenclature for Glycans (SNFG) system.



Supplementary Figure 2.5 Two-dimensional and three-dimensional k-mean cluster of the relative abundance of each cell line (Caco-2, A549, and PNT2) with and without inhibition treatments (Kifunensine, 2,4,7,8,9-Penta-O-acetyl-N-acetyl-3-fluoro-b-D-neuraminic acid methyl ester, 2-deoxy-2-fluorofucose, and 6-alkynyl fucose. Red cluster 1 (untreated control cells), dark green cluster 2 (2FF treated cells), dark blue cluster 3 (3FS treated cells), orange cluster 4 (Kif treated cells) and light green cluster 5 (6AF treated cells). The plot was created with JMP Pro 15 software using K-mean clustering analysis.

Chapter III

**Novel Characterization of the Glycocalyx of 3D Culturing Model
Identified Glycans as Potential Biomarker for Colorectal Cancer**

Abstract

Extensive glycolyx profiling methods were employed to identify the difference between two-dimensional (2D) and three-dimensional (3D) culture models with two human colorectal cancer HCT116 and HT29. Three-dimensional cell cultures have become more commonly used in cancer research due to their ability to mimic the microenvironment found in tumors. Various “omics” analyses (genomic, proteomic, and metabolomic) have been performed on 3D models. However, none have characterized the glycolyx, even though glycosylation plays a critical role in cancer development. This research performed the earliest analysis of N-glycans, O-glycans, glycolipids, proteomic, and glycoproteomic on the cell membranes obtained from 2D and 3D models of HCT116 and HT29 cancer cell lines. The N- and O-glycans were separated with liquid chromatography using a PGC chip column, and the glycolipids were separated with a nanoflow LC-chip with a C18 stationary phase. The peptides and glycopeptides were separated with a C18 nanoflow LC column. The structures were profiled using a nanoliquid chromatography-mass spectrometry (nanoLC–MS). The 2D omic results were compared to the 3D omic results for both cell lines. The analysis of the glycolyx showed a 20% increase in high-mannose glycans, an increase in core fucosylated glycans, and an increase in sialylated glycolipids in the 3D models for both cell lines. Glycoproteomic and proteomic results showed significant upregulation in multiple biological pathways that are related to cancer formation.

Introduction

Colorectal cancer (CRC) is one of the most commonly diagnosed cancers worldwide, and the second most common cause of cancer death in the United States¹. Early treatments of CRC yield a high survival rate of 91%, however early detection of CRC is challenging because it shows little to no symptoms in early development.² Additionally, current detection methods lack the level

of sensitivity and specificity required for detection³. Therefore, additional efforts are necessary to understand better the underlying biological mechanism of CRC to develop a more reliable method of detecting and diagnosing CRC.

Biomarkers for CRC progression include an increase in sialylated and fucosylated glycans expression in the glycocalyx⁴. The glycocalyx is a carbohydrate-rich environment found on the cell membranes composed of an array of glycolipids and glycoproteins -with N- and O- glycans attached. The glycosylations of biomolecules are crucial for regulating several physiological and pathological functions, namely, immune response, protection against infection, and cell-cell interactions. Moreover, aberrant glycosylation plays an essential role in various biological pathways of different diseases such as cancer. Some cancer-associated alterations in the glycocalyx are increased sialylation in proteins and lipids, increased core fucosylated N-glycans, and increased high-mannose glycans⁵⁻⁹.

In vitro cancer models are necessary tools for cancer research as they provide a low-cost platform to investigate human biological mechanisms without the same ethical limitations of working with human tissues. Conventional cancer research is performed on *in vitro* models (a two-dimensional (2D) cell culturing model), followed by *in vivo* models (animal model) to validate the findings. However, the 2D model has limitations as a tumor model because it lacks certain elements found in tumors, namely artificial attachment of the cells, small cellular interaction, and only formation of monolayers. On the other hand, *in vivo* models can mimic those elements, but *in vivo* methods are costly, time-consuming, and hard to reproduce similar results in clinical trials¹⁰⁻¹². Therefore three-dimensional (3D) cell culturing helped bridge the gap between *in vitro* and *in vivo* studies.

A three-dimensional model can mimic the microenvironment and cell-cell interactions that are found in tumors. For those reasons, 3D culturing is becoming more prevalent in cancer research and has been used for drug discovery to combat various types of cancer, e.g., colorectal, breast, and glioblastoma¹³⁻¹⁵. Currently, various omic analyses have been performed to characterize the 3D models, including proteomic, lipidomic, phosphoproteomic, and metabolomic^{16,17}. These profiling methods give vital information about biological functions in the model. Nevertheless, an extensive analysis of the glycocalyx is still lacking for the N-, O-glycome, glycolipidome, and glycoproteome.

General interest in glycosylation has led to recent developments in new methods to characterize the glycocalyx particularly based on nano-Liquid Chromatography-Mass Spectrometry (nanoLC-MS)^{18,19}. The high sensitivity of nanoflow allows the characterization of low abundant glycans, glycolipids, and glycoproteins in many biological systems²⁰. This workflow utilizes nanoflow liquid chromatography-quadrupole time-of-flight MS (nanoLC-QTOF MS) and a nanoflow liquid chromatography orbitrap MS (nanoLC-Orbitrap MS) to perform a comprehensive analysis of the glycocalyx.

In this research, we performed extensive omics analysis on the 2D and 3D culturing models of two human colorectal cell lines (HCT116 and HT29). We monitor the N-glycome, O-glycome, glycolipidome, proteome, and glycoproteome of the 2D and 3D of HCT116 and HT29 and quantitate the difference between the two culturing models for both cell lines. In addition, a spatial N-glycomic profile was performed for the 3D model of both cell lines. The glycocalyx analyses were performed using a nanoLC-chip-QTOF MS, and the proteomic and glycoproteomic analyses were performed using a nanoLC-Orbitrap MS.

Methods and Materials

HCT116 and HT29 cell lines, and McCoy's 5A medium, were obtained from the American Type Culture Collection (ATCC). Dithiothreitol (DTT), iodoacetamide (IAA), and ultra-low attachment 96 well plate were purchased from Sigma-Aldrich. Phosphate Buffered Saline (PBS), fetal bovine serum (FBS), penicillin, and Accutase were purchased from ThermoFisher Scientific. Sequencing grade trypsin and Glu-C were purchased from Promega. ISPE-HILIC cartridges were purchased from HILICON AB (Sweden).

Cell Line Culture

Human colon colorectal adenocarcinoma epithelial HT29 cells and human colon colorectal carcinoma epithelial HCT116 cells were obtained from ATCC and grown in McCoy's 5A medium supplemented with 10% (v/v) FBS, and 1%(v/v) penicillin incubated at 37°C in an atmosphere of 5% CO₂. Media were replaced every 48 h. Ultra-low attachment 96 well plates (Sigma-Aldrich) were used for spheroid culture. Each well was seeded with 5000 cells supplemented with 150 µL of media.

Enzymatic Accutase Peeling of Spheroids

Cells from spheroids were fractionated into three distinct physiological populations using accutase (ThermoFisher Scientific) treatment. Accutase was warmed to 37°C before treatment. Spheroids were washed twice with PBS to remove any media in the cell culture dish, followed by the addition of 2 mL of accutase. The cells were incubated at 37°C and shaken at 70 rpm for 5 mins. The detached cells in accutase were collected (proliferation layer), then washed again with 2 mL of PBS. The process was repeated once more to collect the hypoxia layer. The remaining cells were collected and labeled as the necrotic core. The three fractions were subjugated to

centrifugation and stored in HB buffer at -80°C until prepared for cell membrane extraction. Ninety-six spheroids were used for N-glycomic analysis.

Cell Membrane Extraction

Extraction protocols were described previously and applied here with slight modifications²¹. Briefly, Cellular debris and mitochondrial fractions were pelleted by centrifugation at 2,000 x g for 10 minutes. Supernatants were transferred to perform ultracentrifugation at 200,000 x g for 45 minutes at 4 °C. Pelleted plasma membrane was then resuspended in 500 µL of 0.2 M Na₂CO₃ and 500 µL of nanopure water followed by two series of ultracentrifugation at 200,000 x g for 45 minutes to wash off cytoplasmic and endoplasmic reticulum fractions.

Protein Digestion

Digestion protocols were described previously and applied here with slight modifications²². Cell pellet was sonicated with 60 µL of 8M urea for 20 min followed by the addition of 2 µL of 550 mM DTT incubated at 55°C for 50 min. After incubation, 4 µL of 450 mM of IAA was added and placed in the dark for 20 min. Then, 420 µL of 50 mM ammonium bicarbonate (ABC) was added, followed by 2 µg of trypsin (reconstituted in 50 mM ABC). An additional 2 µg Glu-C (reconstituted in 50 mM ABC) was added for the glycoproteomic sample. All samples were then incubated at 37°C for 18 h. For proteomic analysis, peptides were desalted using C18 cartridges SPE. For glycoproteomic analysis, the samples were dried and desalted using HILIC cartridges SPE. The peptide and glycopeptide concentrations were determined with bicinchoninic acid assay (BCA) peptide assay (ThermoFisher Scientific) prior to injection in the Orbitrap Fusion Lumos nanoLC-MS/MS instrument.

Glycocalyx Extraction

Glycocalyx extraction were described previously and applied here with slight modifications²³. After cell membrane extraction, the membrane pellet was reconstituted in 100 μ L of 100 mM ABC and 5 mM DTT. Each sample was heated at 100°C for 1 min to denature proteins. An aliquot of 2 μ L of 500,000 U/mL peptide N-glycosidase F (PNGase F) was added to each sample and microwaved at 60°C for 10 min to accelerate N-Glycans release. Afterward, the samples were incubated for 18 h at 37°C to hydrolyze the N-Glycans. The reaction was quenched with 350 μ L of water and went under ultracentrifugation at 200,000 x g. The supernatant was extracted and cleaned with porous graphite carbon (PGC) SPE and dried. The remaining pellet was dried under vacuum for 10 min, followed by Folch or Bligh–Dyer extraction to separate the glycolipids. The supernatant was extracted after centrifugation at 8,800g for 5 min at 25 °C. The addition of 100 μ L of 0.1 M potassium chloride was added to the supernatant and centrifuged for 8,800g for 5 min at 25 °C. The aqueous layer was extracted, dried, washed with C8 SPE plate, and dried again. The remaining pellet was dried again under vacuum, then rehydrated with 90 μ L of water and sonication for 20 min. O-glycans were chemically released using B-elimination and were performed by adding 100 μ L of 2 M NaBH₄ and 10 μ L of 2 M NaOH. Samples were incubated at 45 °C for 18 h. Samples were cleaned with PGC SPE, dried, HILIC SPE, and dried again. N-Glycans and O-glycans were reconstituted in water, and glycolipids were reconstituted in 1:1 (vol/vol) methanol/water before NanoLC-chip-Q-TOF MS Analysis.

Glycocalyx nanoLC chip-QTOF Analysis

The glycocalyx analysis was described previously and applied here with slight modifications²³⁻²⁵. The N-glycans, the O-glycans, the glycolipids were separated using a nanoLC Chip–QTOF MS with a 65 min, where the glycans begin to elute between 10 and 40 min. N-

glycans, O-glycans, and glycolipids separation was conducted with a binary solvent system. Solvent A composed of 3% (v/v) acetonitrile (ACN) and 0.1% (v/v) formic acid (FA) in water, and solvent B composed of 90% (v/v) ACN and 1% (v/v) FA in water for N- and O-glycans separation. Solvent A composed 20 mM ammonium acetate and 0.1% acetic acid in water, and solvent B composed 20 mM ammonium acetate and 0.1% acetic acid in 85:15 (v/v) methanol/isopropanol for glycolipid separation. The N- and O-glycans were injected into an Agilent PGC microfluidic chip consisting of a 40 nL enrichment and a 43 mm × 75 μm analytical column, with a particle size of 5 μm. The gradient sequence for the N- and O-glycans analysis was: 0–2.5 min, 1% B; 2.5–20 min, 16% B; 20–35 min, 58% B; 35–40 min, 100% B; 40–50 min, 100% B; 50.01–65 min, 0% B with a flow rate of 0.3 μL/min. The glycolipids were injected into an Agilent C18 microfluidic chip consisting of a 40 nL enrichment and a 150 mm × 75 μm analytical column, with a particle size of 5 μm. The gradient sequence for the glycolipid analysis was: 0–1 min, 70% B; 1–4 min, 85% B; 4–40 min, 100% B; 40–55 min, 100% B; 55–58.1 min, 70% B; 58.1–65 min, 70% B with a flow rate of 0.3 μL/min. The N-glycans, O-glycans, and glycolipids analysis mass range was m/z 600–2000, m/z 400–2000, and m/z 500–2000, respectively, with spectra measured at 0.8 s per spectrum in positive mode ionization. Reference mass m/z 1221.991 was used to correct mass inaccuracies. The cell-surface glycosylation changes were quantified using MassHunter software (Agilent, Santa Clara, CA, USA, version B.07.00).

Proteomic nanoLC-MS/MS Analysis

Proteomic analysis of peptides was separated onto a C18 column, 150 mm × 75 μm with a particle size of 2 μm 100 Å, at a flow rate of 0.3 μL/min with a binary solvent system with solvent A containing 0.08% (v/v) formic acid (FA) in water and solvent B containing 80% ACN (v/v) 0.1% (v/v) FA in water. Solvent gradient for peptide separation was: 0 min 4% B; 0–3 min, 9% B; 3–6

min, 12% B; 6-80 min, 24% B; 80-95 min, 50% B; 95-99 min, 99% B; 99-103 min, 99% B; 103-103.5 2% B 103.5-120 min 2%B. Mass range of m/z 375-1500 for MS1 and mass range of MS2 was set to auto. Raw output files were then inputted into Byonic software for peptide analysis, followed by an extracted ion chromatogram (EIC) using Byologic software (Protein Metrics).

Glycoproteomic nanoLC-MS/MS Analysis

Glycoproteomic analysis was described previously and applied here with slight modifications²³. Glycopeptides were separated onto a C18 column, 150 mm × 75 μm with particle size of 2 μm 100 Å, at a flow rate of 0.3 μL/min with a binary solvent system with solvent A containing 0.08% (v/v) FA in water and solvent B containing 80% ACN (v/v) 0.1% (v/v) FA in water. Mass range of m/z 700-2000 for MS1 and mass range of MS2 was m/z 120 and greater. Raw output files were then inputted into Byonic software for glycopeptide analysis followed by EIC using Byologic software (Protein Metrics).

Apoptosis Assay

Three-dimensional spheroids of HCT116 and HT29 cells were cultured at 37 °C, 5% CO₂ for 14 days. First, spheroids were collected from ultralow-attachment plates by pipetting, and then the proliferation, hypoxia, and necrotic core layers were separated using accutase treatment. The apoptosis assay was performed using Nexcelom Annexin V-FITC/PI™ Apoptosis kit according to manufacturer instructions. After harvesting the cells via trypsinization, aliquots were prepared to contain approximately 600,000 cells for each sample. The cell suspensions were subsequently centrifuged at 400 g for 5 minutes and then resuspended in 40 μL Annexin V binding buffer, making sure to resuspend the cells completely. Then, 5 μL each of Annexin V-FITC and PI stains were added to the cell suspensions. These were incubated for 15 minutes at 37°C in the dark. After incubation, 250 μL of PBS was added to each sample and then centrifuged at 400 g for 5 minutes,

carefully removing the supernatant. The cells were finally resuspended in 100 μ L Annexin V binding buffer before data acquisition. Data was acquired using Cellometer Vision CBA 5, using the protocol CBA_Annexin V+PI assay, with an F1 exposure time of 8000 msec and F2 exposure time of 20000 msec. Acquired data were analyzed using FCS Express 7.0. Cell gating was adjusted using the proliferation layer cells as control. Assays were done in triplicate.

Quantitative Proteomics and Glycoproteomics Data Analysis

The mass spectrometry data were analyzed using Byos workflow (Protein Metrics). For qualitative analysis in Byonic (Protein Metrics), proteins were compared to the human proteome database using a precursor mass tolerance of 20 ppm and fragment mass tolerance of 10 ppm²⁶. The digestion parameters used included C-terminal cleavage by trypsin (K and R cleavage sites) with at most two missed cleavages. The following peptide modifications were included: carbamidomethyl @ C, oxidation @ M, deamidation @ N and Q, acetylation at protein N-terminal, Gln to pyro-Glu at N-terminal Q, Glu to pyro-Glu at N-terminal E. Protein IDs were filtered at 1% FDR. To identify the glycoproteins and glycoforms, an additional search was performed in Byonic using an in-house N-glycan database. Quantification for each protein was done in Byologic (Protein Metrics) by quantifying the XIC area sum of the top 3 most abundant peptides. XICs were then normalized to sum total before statistical analysis. On the other hand, glycoform quantification was normalized to each protein's glycosite to yield the percentage occupancy of a particular glycoform.

Gene Ontology Analysis of Proteomics and Glycoproteomics

To identify significantly different proteins and glycopeptides, multiple t-tests were conducted in GraphPad Prism using an FDR approach (FDR=5%). Significantly over- and under-expressed proteins were annotated using g:profiler software and STRING^{27,28}. Comparative

protein expression between 2D- and 3D-cultured HT29 and HCT116 cells were mapped out using Metaboanalyst software²⁹. Similarly, the glycopeptides were annotated using STRING to yield significantly enriched KEGG pathways. Significantly different glycopeptides were plotted as a heatmap in Morpheus (Broad Institute. (n.d.). Morpheus. <https://software.broadinstitute.org/morpheus/>).

Glycoproteomic Profiling of ENPL and ITGA6

Based on the glycoproteomics analysis, ENPL (P14625) and ITGA3 (P26006) were identified as potentially significant glycoproteins in HT29 and HCT116, respectively. The glycoforms quantified from these proteins were extracted and compared between 2D- and 3D-cultured cells. The glycoforms that had the greatest difference in abundance were selected to be modelled into ENPL and ITGA3. Protein structures of ENPL and ITGA3 were homology modelled using SWISS-MODELLER (<https://swissmodel.expasy.org/>)³⁰ using GRP94 (PDB ID: 2o1u) and ITGAV (PDB ID: 3ije) as templates for ENPL and ITGA3, respectively. PDB files of these homology models were subsequently glycosylated at specific asparagine residues (obtained from glycoproteomics result) using Glycan-modeller in CHARMM-GUI (<https://charmm-gui.org/?doc=input/glycan>)³¹. The modelled fully-glycosylated ENPL and ITGA3 were then assessed for glycan-protein using the “Find clashes/contacts” tool in Chimera, which considers atoms that have a VDW overlap $\geq -0.4 \text{ \AA}$ ³².

Results

An extensive glycomic workflow for comparative characterization of the 2D and 3D models is shown schematically in **Figure 3.1**. It involves cell membrane enrichment of both models subjected to N-glycomics, O-glycomics, glycolipidomics, proteomics, and glycoproteomics analysis of two colon cancer cell lines (HCT116 and HT29). The N-glycome, O-

glycome, and glycolipidome were characterized with a nanoLC-MS with a QTOF mass analyzer. The N- and O-glycome analysis used a PGC stationary phase, while the glycolipidome used a C18 stationary phase for separation. The proteome and glycoproteome were characterized with a nanoLC-MS using a C18 nanocolumn and Orbitrap mass analyzer to produce their profiles. This workflow has been reported in previous publications²³. We identified over 300 N-glycans, 100 O-glycans, and 100 glycolipids structures -including isomers. In addition, around 1000 proteins were identified in both cell lines from the proteomic analysis. Over 600 glycopeptides were identified from HCT116, and over 900 glycopeptides were identified from HT29 cells.

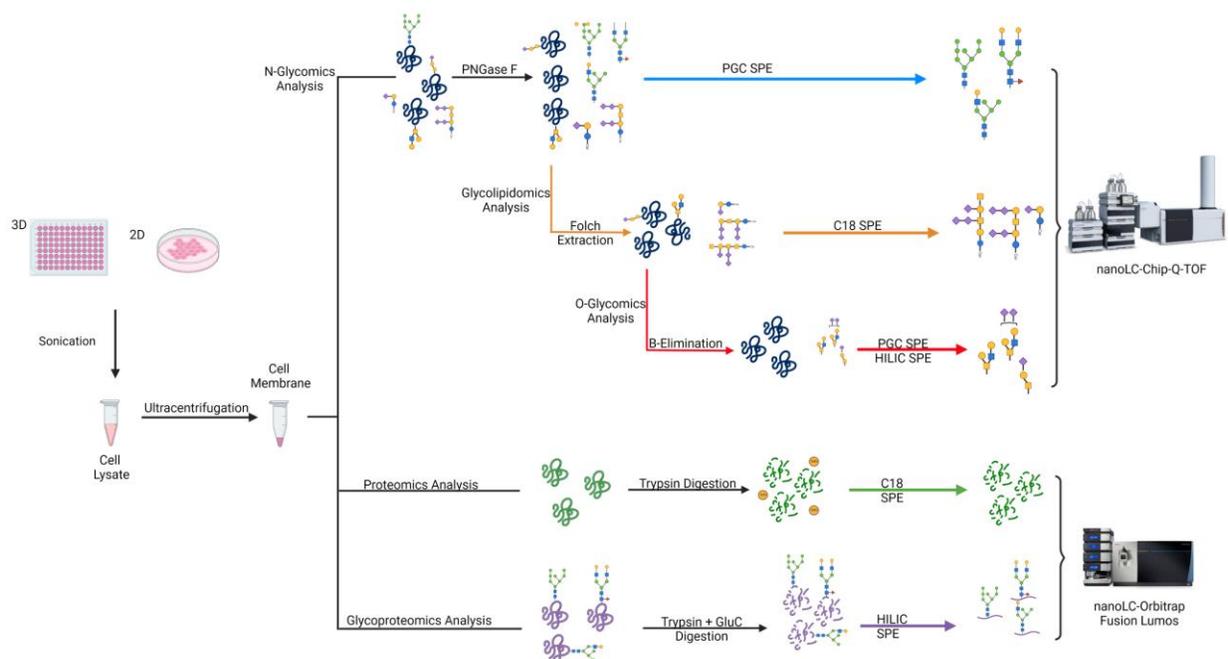


Figure 3.1 The schematic workflow for the characterization of the glyocalyx analysis of the 2D and 3D cell culture model with HCT116 and HT29 cells. The 2D and 3D samples were collected and separated into three groups: the glycome analysis of the glyocalyx, the proteomic analysis, and the glycoproteomic analysis.

N-Glycomic Profile of 2D and 3D Cell Membrane

The N-glycomic profiles for 2D and 3D models of HCT116 (colon epithelial carcinoma cell line) and HT29 (colon epithelial adenocarcinoma cell line) were determined, and each yielded over 300 N-glycans (including isomers) **Figures 3.2A** and **3.2B**. N-Glycans were conventionally categorized into three types, high-mannose-, hybrid-, and complex-type glycans. The N-glycome profile was drastically altered when the cell lines were cultured in the 3D model. There was a two-fold increase in high-mannose N-glycans and a decrease in complex-type N-glycans in both 3D models compared to their 2D counterparts. We expanded the categories further by grouping the N-glycans into high-mannose, undecorated (nonfucosylated nonsialylated), fucosylated (only), sialylated (only), and sialofucosylated to provide more specific comparisons of the structures (**Figure 3.2C**). Both cell lines showed significant changes in their N-glycome profile with q-values no more than 0.005 based on a multi-T-test analysis with 1% FDR correction. The N-glycome profile of HCT116 showed a significant increase in high-mannose N-glycans, from 23 percent in the 2D model to 43 percent in the 3D model, along with major decreases in undecorated (7 percent) and sialylated (12 percent) N-glycans between the 2D the 3D cells. HT29 showed similar behavior with a 20 percent increase in high-mannose N-glycans and a 13 percent reduction of sialylated N-glycans. In addition, HT29 significantly increased in undecorated (11 percent) and fucosylated (13 percent) N-glycans and decreased in sialofucosylated (31 percent) N-glycans. It is noted that HCT116 lacked fucosylated and sialofucosylated N-glycans. The lack of fucose is expected due to a mutation of the GMDS gene. This gene encodes for GDP-mannose 4,6, dehydratase protein, a crucial enzyme for the synthetic pathway of fucose⁴. When looking at the individual N-glycans, HCT116 had 19 significantly altered N-glycans, and HT29 had nearly 25 altered N-glycans with statistical significance as shown in **Figure 3.2D** (with q-value lower than 0.003). In HCT116, there were twelve N-glycans found to be up-regulated in the 3D model- seven of them being high-

mannose glycans. The remaining N-glycans were down-regulated in 3D models, with the most abundant N-glycan being Hex₆HexNAc₅. HT29 showed a similar trend to HCT116, where about 80% (17 N-glycans) of the significantly altered N-glycans were up-regulated glycans in the 3D model. This fraction is composed of N-glycans, either undecorated or fucosylated, and only one high-mannose glycans (Hex₃HexNAc₂). If we removed the FDR correction from HT29 multi-T-test analysis, all the high-mannose N-glycans in HT29 are statistically significant with p-values less than 0.05 when comparing the 3D model to the 2D model. In addition, the increase of fucosylated glycans was mainly singly fucosylated glycans, where the fucose are typically in the core structure.

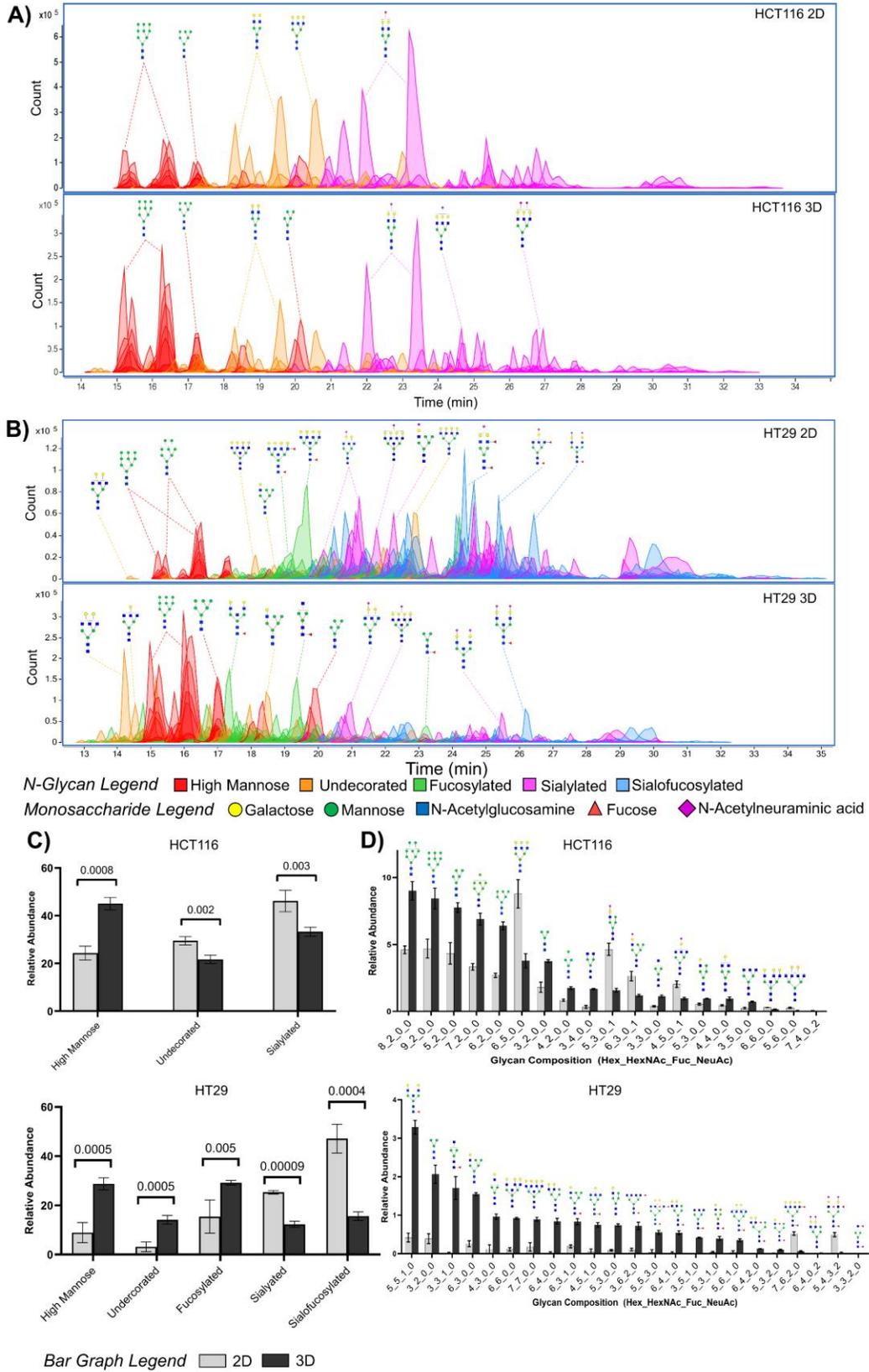


Figure 3.2 Extracted compound chromatograms of the N-glycan profile of the 2D and 3D culture model where (A) are HCT116, (B) is HT29, (C) the significant glycan categories top HCT116 and bottom HT29 with their q values, and (D) is all the significantly altered top HCT116 and bottom HT29. The peaks are colored by glycan subtypes and annotated with the schematic representation of the glycan structures. Multi-T test was performed with 1% FDR correction.

Spatial N-Glycomic Profile of 3D Cell Membrane

To further study the complex system of 3D culturing models, an enzymatic peeling of the spheroid was performed with accutase. The peeling assay fractionated the spheroids into three fractions, the outer layer (proliferation layer), the middle layer (quiescent layer), and the inner layer (necrotic core). The N-glycomic profile of each layer (for both cell lines) was determined, and each yielded at least 150 or more N-glycans (including isomer) in **Figures 3.3A** and **3.3B**. The high-mannose glycans were the most abundant in the proliferation layer (41 percent in HCT116 and 32 percent in HT29) of both cell lines. Additionally, the necrotic core yielded more N-glycans than the proliferation layer, composed of more complex-type N-glycans. The relative abundance of each N-glycans was plotted for both cell lines in **Supplementary Figures 3.1** and **3.2**. An apoptosis assay was performed on the three fractions isolated from the peeling assay for both cell lines (**Supplementary 3.3C** and **3.3D**). The final fraction (the necrotic core) in both cells showed a significant increase in dead cells compared to the proliferation and quiescent layer, especially for HT29, where the percentage of live cells dropped from the high nineties to less than sixty percent. In addition, a recording (**Supplementary Video 3.1**) of the accutase peeling of HCT116 was taken. The enzymatic peeling video of HCT116 did not show any disruption to the 3D

structure of the spheroid. Additional omic-analysis was not performed on the three fractions due to limited samples from the enzymatic peeling.

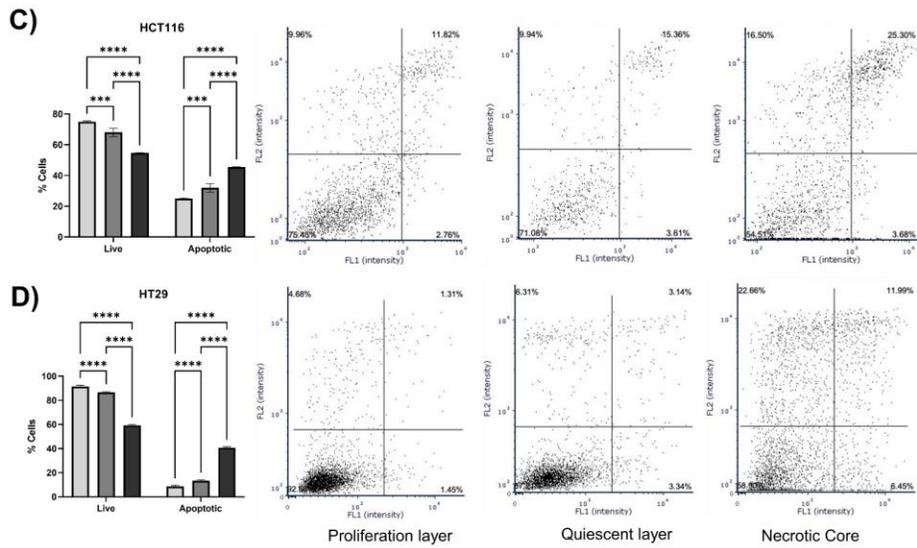
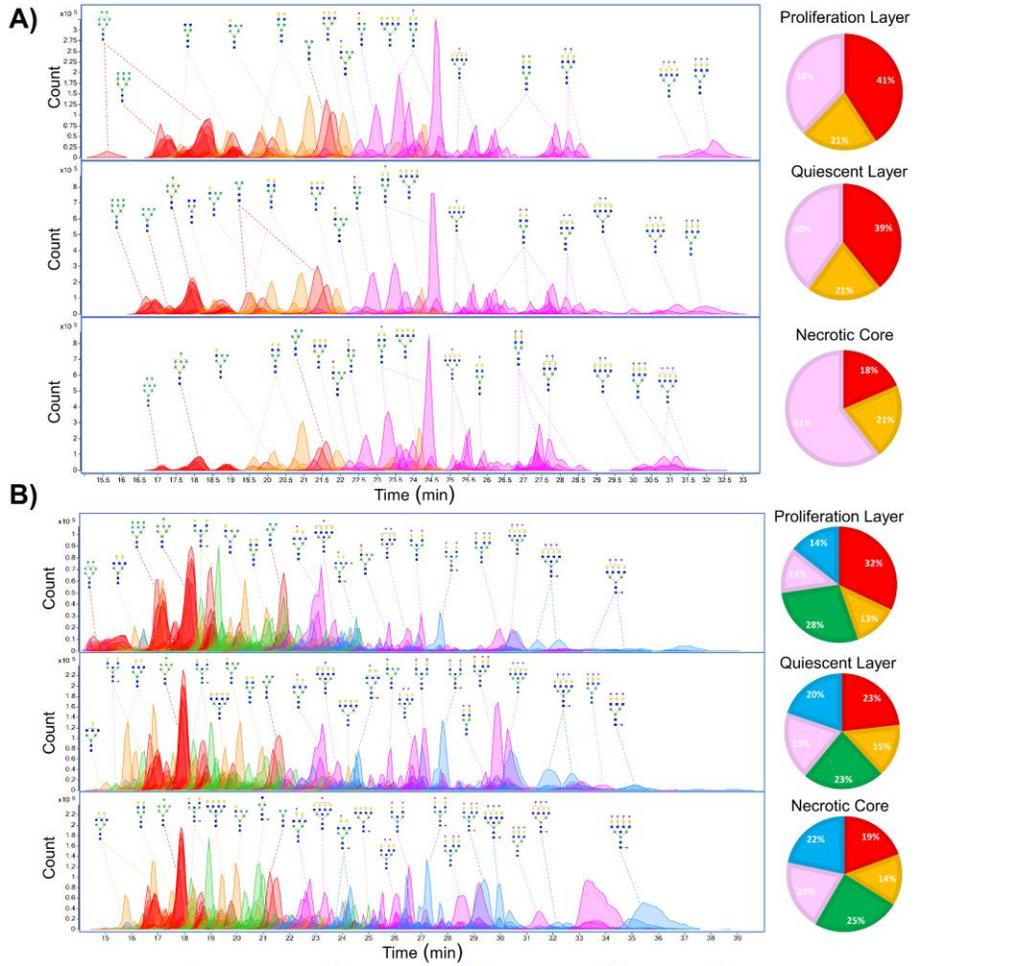


Figure 3.3 Extracted compound chromatograms of the spatial profile of N-glycans where (A) is HCT116, (B) is HT2, (C) apoptosis assay of HCT116 with the significant change and (D) apoptosis assay of HT29 with the significant change in dead cells. The peaks are colored by glycan subtypes and annotated with the schematic representation of the glycan structures. Two-tailed t-test was performed with p-value < 0.05 is *, < 0.01 is **, < 0.001 is ***, and < 0.0001 is ****.

O-Glycomic Profile of 2D and 3D Cell Membrane

The O-glycans were released using beta-elimination after glycolipid extraction, and the O-glycan structures were determined, and each yielded around 100 O-glycans (including isomers) **Supplementary 3.4A** and **3.4B**. Unlike the N-glycan profile, the O-glycan profile showed subtle changes. The 3D profile compared to the 2D profile of HCT116 had a more noticeable difference in the chromatogram than HT29. A consistent trend in both cell lines was an increase in sialylated O-glycans in the 3D culturing model **Figure 3.4C**. With individual O-glycans, the 3D model of HCT116 had a significantly decreased O-glycan (Hex₄HexNAc₄) and HT29 had a significantly increased O-glycan (Hex₁HexNAc₁Fuc₁NeuAc₁) (**Figure 3.4D**). A common trend is found in both cell lines, with Hex₂HexNAc₂NeuAc₂, Hex₂HexNAc₁NeuAc₂, and Hex₁HexNAc₁NeuAc₁ increasing in the relative abundance in the 3D model compared to the 2D.

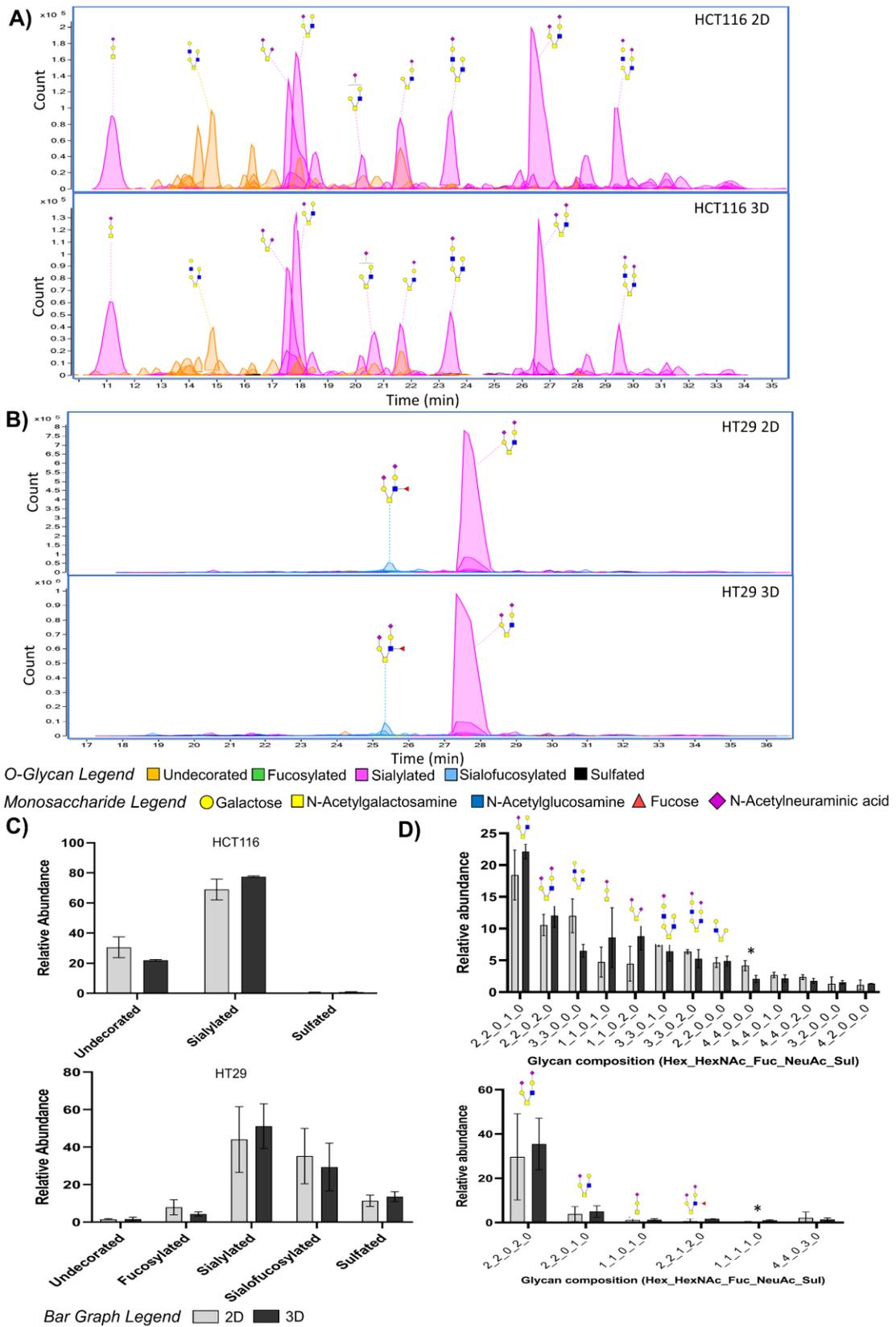


Figure 3.4 Extracted compound chromatograms of the O-glycan profile of the 2D and 3D models where (A) is HCT116, (B) is HT29, (C) O-glycan categories of HCT116, and (D) O-glycan categories of HT29. The peaks are colored by glycan subtypes and annotated with the schematic representation of the glycan structures. Two-tailed t-test was performed with p-value < 0.05 is *.

Glycoproteomic Analysis of 2D and 3D Cell Membrane

Site-specific glycoproteomic analysis was performed on 2D and 3D culturing models for both cell lines. The resulting data were plotted on a volcano plot shown in **Figure 3.5A**. There were over 950 quantifiable N-glycopeptides (including glycoforms) identified in HT29, representing 170 glycoproteins. Among those, 250 N-glycopeptides were found to be significantly altered, representing 68 N-glycoproteins. HCT 116 had over 600 quantifiable N-glycopeptides (including glycoforms) identified corresponding to 116 N-glycoproteins. Of those, 135 N-glycopeptides were considered significantly altered, representing 44 glycoproteins. The significantly altered N-glycopeptides were aggregated into five categories: high-mannose, undecorated, fucosylated, sialylated, and sialofucosylated, and these glycopeptides were plotted in a heatmap shown in **Figure 3.5B**. HT29 has higher N-glycopeptides abundance in the 3D model, though there was no noticeable trend on which specific N-glycans categories were favored. On the other hand, HCT116 had higher N-glycopeptides abundance in the 2D model.

Gene ontology and STRING network analysis were performed on the significantly altered N-glycopeptides (**Figure 3.5C**) to study the protein interactions and pathways affected by the difference in the 3D models to the 2D counterpart. Both HT29 and HCT116 have pathways in proteoglycans in cancer, cell-adhesion molecules, focal adhesion, ECM-receptors interaction, P13K-Akt signaling, phagosome, and proteins processing in the endoplasmic reticulum that were found to be significantly affected. In addition, HT29 has significantly altered pathways in cancer

and N-glycan biosynthesis. Using the information, we performed homology modeling on a significantly altered glycoprotein from each cell line (**Figure 3.5D**). ENPL was an example glycoprotein from the HT29 cell line. In the 3D model, ENPL's N-glycans are spread out, while the 2D model N-glycans have a more compact center than the 3D counterpart. This is seen in ASN445, where the 2D model has 34 contacts while the 3D model has 74 contacts. ITA3 was the example glycoprotein from the HCT116 cell line. There were three different amino acid positions that were examined (ASN500, ASN573, and ASN697). In the 2D model, 48, 85, and 44 points of contact in amino acid position ASN500, ASN573, and ASN697, respectively. The 3D model had 52, 63, and 47 points of contact in amino acid ASN500, ASN573, and ASN697, respectively. Interestingly, the HCT116 2D model had more points of contact than the 3D model, which was opposite of the HT29 cell line.

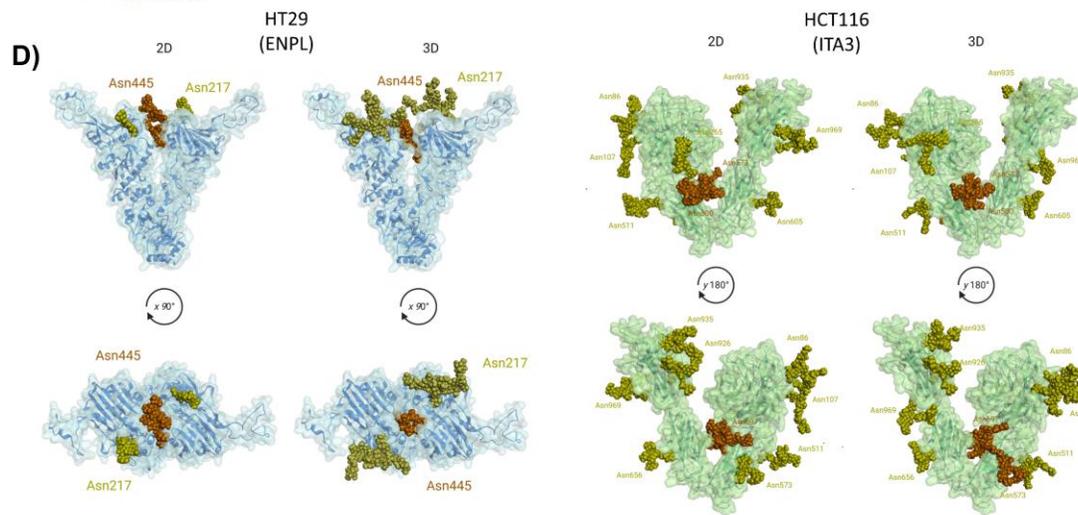
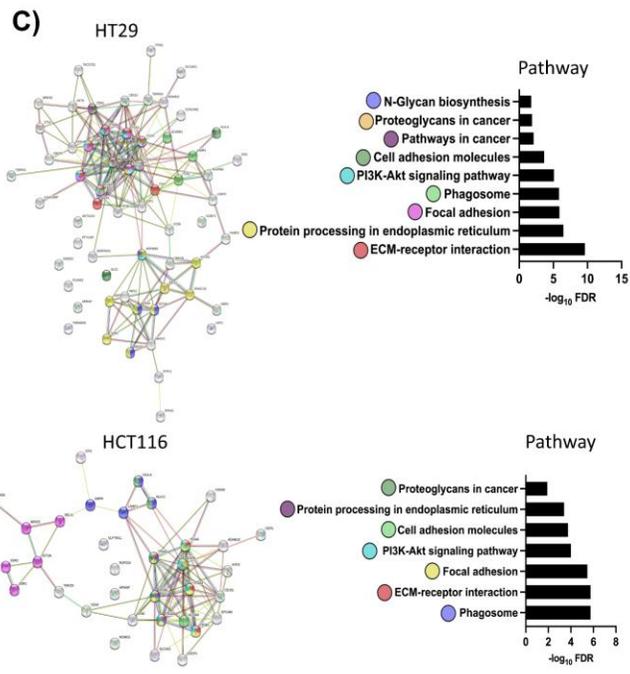
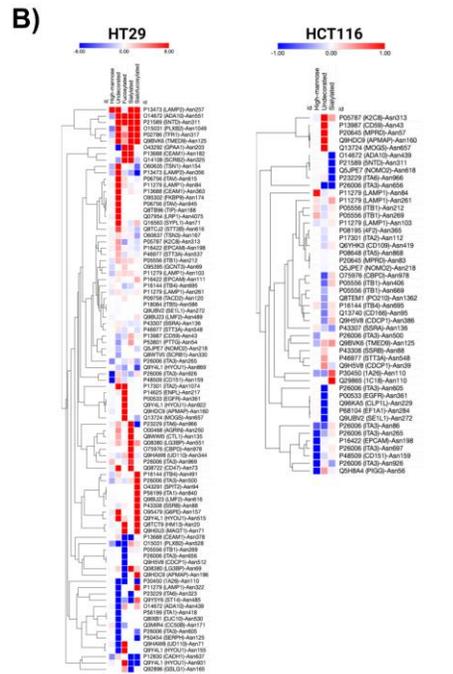
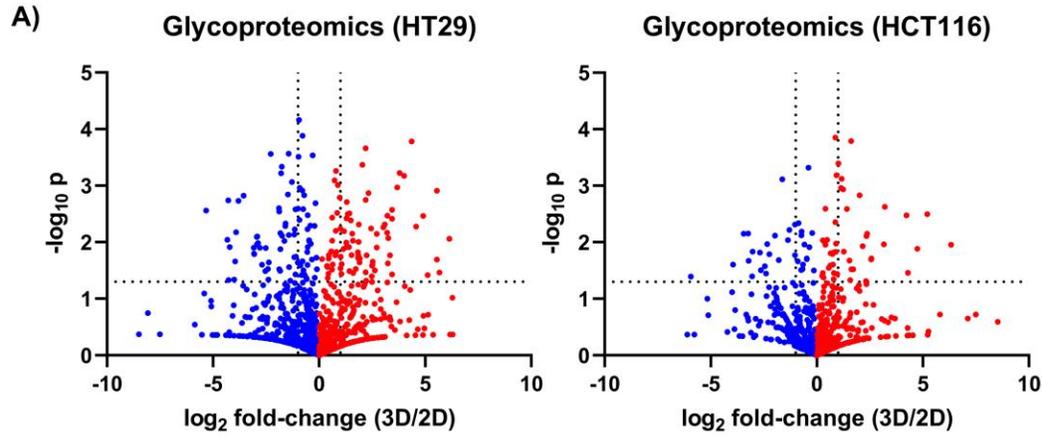
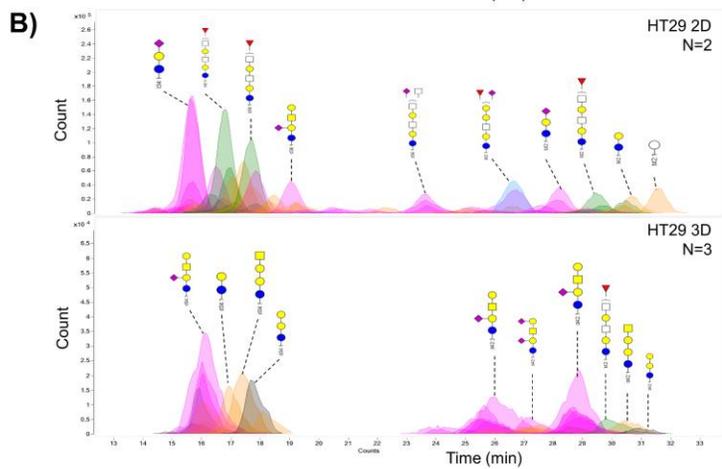
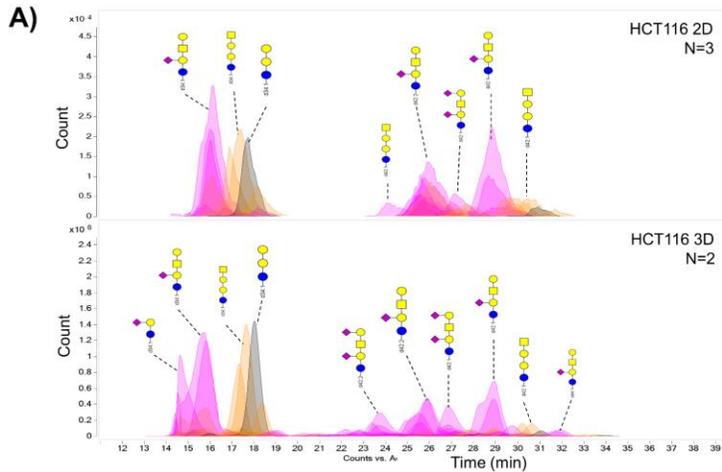


Figure 3.5 Glycoproteomic results of the 2D and 3D model where (A) is the volcano plot of HT29 and HCT116, (B) the heatmap of significant glycopeptides, (C) the Gene ontology and STRING network analysis of significant glycopeptides, and (D) Glycoproteomic simulation of ENPL and ITA3.

Glycolipidomic Profile of 2D and 3D Cell Membrane

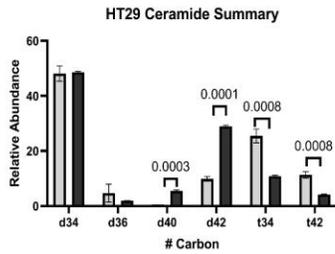
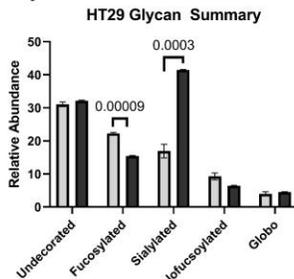
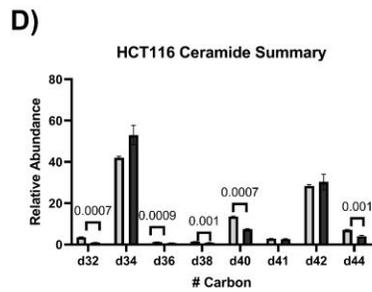
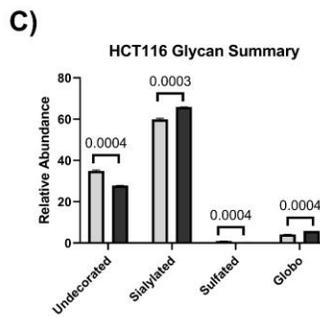
The glycolipids were isolated using Folch extraction after N-glycan release, and the isolated glycolipids were characterized for both cell lines, and each yielded over 100 glycolipids (including isomers) (**Figures 3.6A** and **3.6B**). Both cell lines showed a similar trend, increasing sialylated glycans in 3D models compared to 2D models (5 percent in HCT116 and 25 percent in HT29). Additionally, 2D models seem to express more undecorated (6 percent) and sulfated (0.5 percent) glycans and increase in global (1 percent) glycans in 3D models in HCT116 (**Figure 3.6C**). In HT29, there are more fucosylated (7 percent) glycolipids in the 2D models. These alterations were found to be statistically significant with q values of 0.0003 or less. With the ceramide backbone, the HCT116 model showed more diversity when compared to HT29 (**Figure 3.6D**). In HCT116, the significant ceramide backbone was d32, d36, d38, d40, and d44, with a q value of 0.001 or less. In addition, HT29 decreased relative abundance in t34 and t42 ceramide (which was not found in HCT116) and increased d40 and d42 ceramide in the spheroid model. Both cell lines showed consistency of having d34 as the most abundant ceramide and showed a slight elevation in d34 and d32 in the 3D models (though the alteration was not significant). When looking at the individual glycolipids, 34 glycolipids (including isomer) in HCT116 and 73 glycolipids (including isomer) in HT29 were found to be significant with a 1% FDR correction. Among those, 11 glycolipids were found in both models, namely Hex₂NeuAc₁-d34_1, Hex₃HexNAc₁NueGc₂-d42_1, Hex₃HexNAc₁NeuAc₁-d34_1, Hex₂HexNAc₁-d34_1, Hex₃HexNAc₁-d34_1,

Hex₃HexNAc₁NeuAc₂-d42_1, Hex₂HexNAc₁NueGc₁-d34_1, Hex₂-d34_1, Hex₃-d34_1, Hex₄HexNAc₂-d34_1, Hex₃HexNAc₃NeuAc₁-d40_1. In HT29, one of the largest changed glycolipids was Hex₄NeuAc₃-d34_1 with a log₂Fold change of 6.1 and a q-value of 0.003.



Glycolipid Legend Undecorated Fucosylated Sialylated Sialofucosylated Sulfated Globo

Monosaccharide Legend Galactose Mannose N-Acetylglucosamine Fucose N-Acetylneuraminic acid



Bar Graph Legend 2D 3D

Figure 3.6 Extracted compound chromatograms of the glycolipid profile of the 2D and 3D models where (A) is HCT116, (B) is HT29, (C) the significant glycan categories top HCT116 and bottom HT29 with their q values, and (D) the significant ceramide categories top HCT116 and bottom HT29 with their q values. The peaks are colored by glycan subtypes and annotated with the schematic representation of the glycan structures. Multi-T test was performed with 1% FDR correction.

Proteomic analysis of 2D and 3D cell membrane

Proteomic analysis was performed on the 2D and 3D models in conjunction with the glycocalyx analysis. Like the N-glycan analysis, the proteomic profile showed a drastic change in both cell lines. The volcano plot of the proteomic data is shown in **Figure 3.7A**. HT29 had over 1200 reproducibly quantified proteins, with almost fifty percent of the proteins were significantly altered. HCT116 had over 900 reproducibly quantified proteins, and over thirty percent of the proteins were significantly altered. The proteomic analysis of HT29 is comparable to the previous study of whole-cell proteomic analysis of HT29 2D and 3D culturing¹⁶. There were many upregulation proteins responsible for programmed cell death found in HT29 (**Supplementary Figure 3.3A**). HCT116 had some proteins related to cell death, but not as many as HT29 (**Supplementary Figure 3.3B**). Other biological functions altered between the 3D and 2D models were adhesion, cellular response, and N-glycosylation. Both HT29 and HCT116 had significantly changed proteins in biological processes responsible for central carbon metabolism in cancer (**Supplementary Figure 3.3C**). In addition, HT29 had significantly up-regulated proteins essential in colorectal cancer (**Supplementary Figure 3.3D**). Gene ontology analyses were performed using the significant proteins data from the proteomic analysis. In HT29, 340 pathways

were up-regulated, and 338 pathways were down-regulated. HCT116 had less, 26 were up-regulated, and 113 were down-regulated (**Supplementary Figure 3.4**).

To further expand on the proteomic result, we performed a comparative analysis on all cell lines with their different culturing method. PCA analysis of these data showed that each cell line and its culturing methods were tightly clustered, forming four distinct clusters (**Figure 3.7B**). In addition to the PCA analysis, a heat map was generated with all the data shown in **Figure 3.7C**. Most of the protein expression was unique except for one patch where there was a consistent trend of down-regulated proteins in the 2D models and up-regulated proteins in the 3D models regardless of the cell line. The affected pathway was plotted, displaying the up-regulated or down-regulated pathways based on the proteins' expression level (**Figure 3.7D**). A few pathways were up-regulated in the 3D models that correspond to carbohydrates regulation, namely glucose metabolism, gluconeogenesis, metabolism of carbohydrates, and glycolysis. In addition, some pathways were found to be consistently down-regulated in both 3D models, namely stabilization of p53 and DNA damage checkpoint.

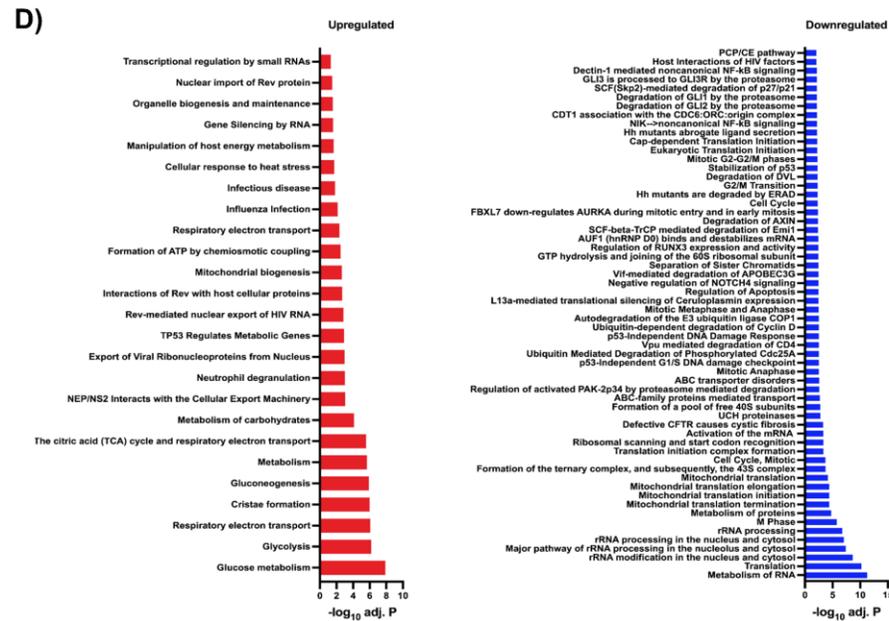
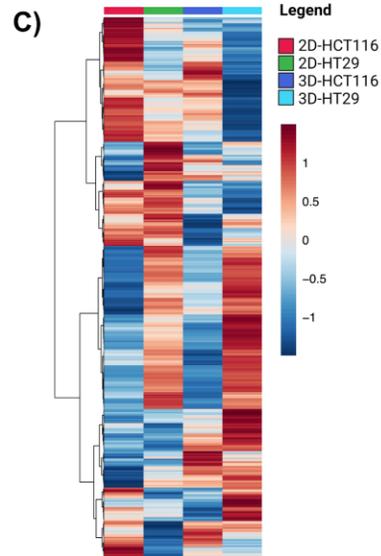
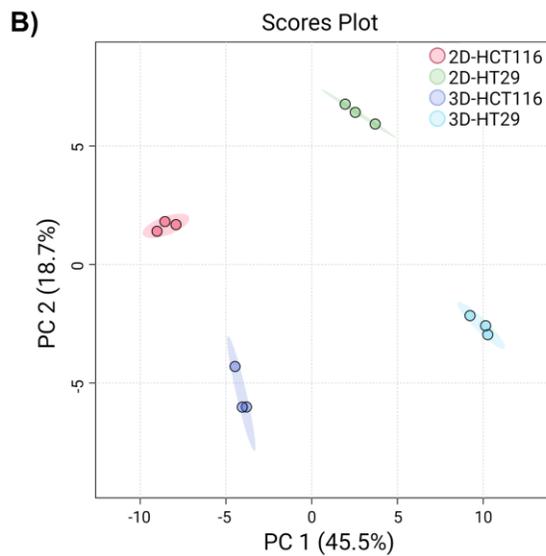
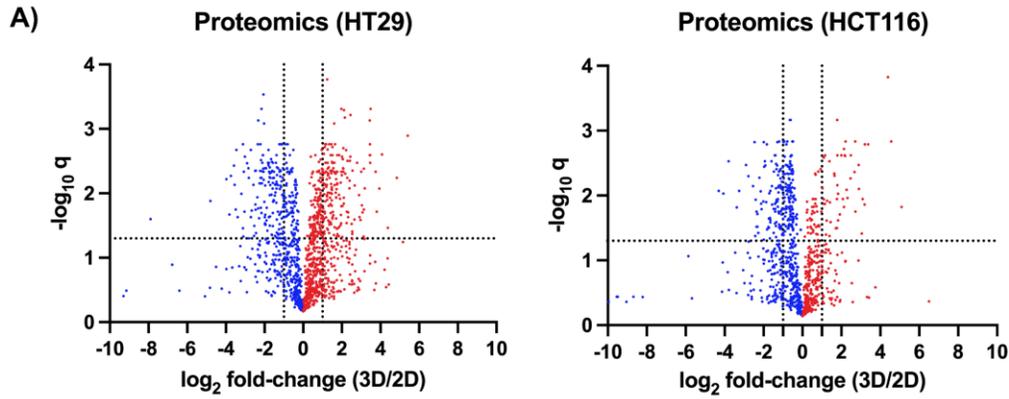


Figure 3.7 Proteomic results of the 2D and 3D model where (A) is the volcano plot of HT29 and HCT116, (B) PCA plot of the 2D and 3D model of both cell line HT29 and HCT116, (C) Heatmap of the 2D and 3D model of both cell line HT29 and HCT116, and (D) The up-regulated and down-regulated proteins that were consistent with the 2D model of HCT116 and HT29 and 3D model. Multi-T test was performed with 1% FDR correction.

Discussions

The N-Glycomic, O-glycomic, glycolipidoimic, proteomic, and glycoproteomic profiles extensively characterize the 2D and 3D culturing models of two types of colon cancer cells (HCT116 and HT29). HCT116 (a microsatellite instability human colorectal carcinoma with a KRAS mutation) and HT29 (a microsatellite stable human colorectal adenocarcinoma with a P53 mutation) represent some of the most common types of mutation found in colorectal cancer^{33,34}. The omic results showed dramatic alteration in both cell lines when comparing the 2D and 3D culturing methods. The alteration showed an upregulation in cancer-related glycans, proteins, and glycoproteins in the 3D model.

The 3D models of HT29 showed an increase in fucosylated glycan (mainly singly fucosylated) compared to the 2D counterpart. No fucosylated glycans were found in HCT116 in this study. Studies have shown the upregulation in fucosylated and sialylated glycans in colorectal cancer cells compared to the non-cancerous cells⁴. Though the N-glycan profile showed a decrease in sialylated glycans, the glycolipid profile showed an increase in sialylated glycans in both models. Thus, the increase of sialylated glycans could also result from an increase in sialylated glycolipids. Additionally, studies have shown the increase in sialylated glycolipid in the cancer tissues cell compared to the non-cancerous counterpart³⁵. Alternatively, both cell lines showed a consistent 20 percent increase in high-mannose type N-glycans. High-mannose glycans could potentially play

an essential role in metastatic characteristics in cancer cells. Past studies showed an increase in high-mannose glycans on breast cancer cell lines made the cells more metastatic by increasing cell migration and cell proliferation rates^{7, 9, 36}. There have been studies on upregulation in O-glycans in colon cancer^{37, 38}. However, the 2D and 3D cell culturing models showed little to no difference.

This study represents the first time the spatial N-glycome profile of spheroid has been reported. Furthermore, the N-glycan profile of both cell lines showed similar trends. The spheroids' most outer layer (proliferation layer) expressed the highest relative abundance of high-mannose glycans in both cell lines. As we go deeper into the non-proliferating layers of the spheroid, the N-glycans composition becomes more diverse with an increasing abundance of complex-type N-glycans. For cell proliferation, it could be essential for cell surface N-glycans to express more high-mannose type N-glycans to allow cells to proliferate. It has been shown that high-mannose N-glycans have faster turnover rates than the complex-type N-glycans³⁹. Therefore, cell proliferation could result in the cells expressing high-mannose N-glycans. The longer the cells do not proliferate, the more opportunity the N-glycans can be added with more diverse monosaccharide moieties (such as fucose), creating more complex-type N-glycans similarly to what is seen in the necrotic core. This could be the cause of why the N-glycome profiles of each layer are different from one another.

STRING analysis of the significantly altered glycopeptide in HT29 and HCT116 indicates a potential effect on pathways affiliated with cancer when comparing the 2D model to the 3D model. These glycoproteins were mainly integrin proteins (ITGAV, ITGA3, ITGA2, ITGB1, ITGA6) and EGFR. For example, ITGB1 at site Asn212 showed a positive increase in undecorated, fucosylated, sialylated, and sialofucosylated glycans in the HT29 cell model. In EGFR, there was

an eight-fold increase in fucosylated glycans at Asn361. However, the proteomic analysis of these two proteins did not show a significant change in the protein expression indicating the effect on the pathway is potentially glycosylation dependent. Another interesting glycoprotein found in HT29 was carcinoembryonic antigen (CEA), a glycoprotein associated with colorectal cancer. This matches a previous study that showed an increase in fucosylated glycans in tumor-associated CEA^{3, 40}. Within the 3D model of HT29, the fucosylated glycopeptide of CEA at Asn363(Hex6HexNAc6Fuc2) had a log₂fold change of 5.5 with a -log₁₀P-value of 4. In addition to the pathway analysis, homology modeling of ENPL and ITA3 were performed. The model demonstrated that the protein to glycan interaction changed when comparing the 2D model to the 3D model. However, the protein conformation stayed relatively the same regardless of what glycans were attached. These increase and decrease in interaction can potentially affect protein functions and communication with other proteins.

Comparing the 2D and 3D proteomic results of both cell lines showed dramatic changes, including the upregulation in cell death and apoptosis pathway. These results suggest the formation of necrotic core in both cell lines' 3D model. In addition, both cell lines display an upregulation in LDHA and KP YM, which is responsible for central carbon metabolism in cancer^{41, 42}. HT29 also had a significant alteration with six proteins (MP2K1, RHOA, RALB, RAC1, CTNB1, KRAS, CYC, and RALA) that were directly correlated to colorectal cancer. These further indicate that the 3D model is more cancer-like which is caused by the increase of cell-to-cell interaction and microenvironment similar to tumors. However, HCT116 did not have up-regulated proteins correlating with colorectal. This could likely be due to the different mutation types between the two cell lines. PCA analysis did show individual clustering between each culturing model and each cell line. The cause is likely due to the difference in mutation since HCT116 is a microsatellite-

instability cell line with KRAS mutation, while HT29 is a microsatellite-stable cell line with a P53 mutation. Additionally, the HT29 3D model had more characteristic traits of upregulation in glycan and protein expression, which could be predominantly due to HT29 being the more prevalent type of colorectal cancer found in humans. These results indicate that the type of mutation and culturing can significantly affect the expression of glycans and proteins found on colorectal cancer, suggesting different mutations or types of cancer may need its own specified treatments.

Conclusion

Herein we reported for the first time the glycocalyx profiles of the 3D culturing model of two colorectal cancer cell lines, HCT116 (KRAS mutation) and HT29 (P53 mutation). Additionally, we reported the N-glycan profiles of the three layers of a 3D spheroid model for the first time. We reported a significant difference with 1% FDR correction between the 2D and 3D models for both cell lines. The resulting values had consistent upregulation in glycans, proteins, and glycoproteins expression that are commonly found in human colorectal cancer. Additionally, we discovered a new trend in the N-glycan profiles suggesting the increase of high-mannose glycans as a potential trait of colorectal cancer formation. We profiled and identified significant alterations in the proteomic and glycoproteomic in the 3D model for both cell lines. The coupling of the glycan composition to the site-specific glycopeptide analysis can provide targeted information to perform protein simulation in order to better understand colorectal cancer. Future studies in colorectal cancer research can compare the *in vivo* models of the glycocalyx on cancer and non-cancerous to further understand the alteration during cancer formation.

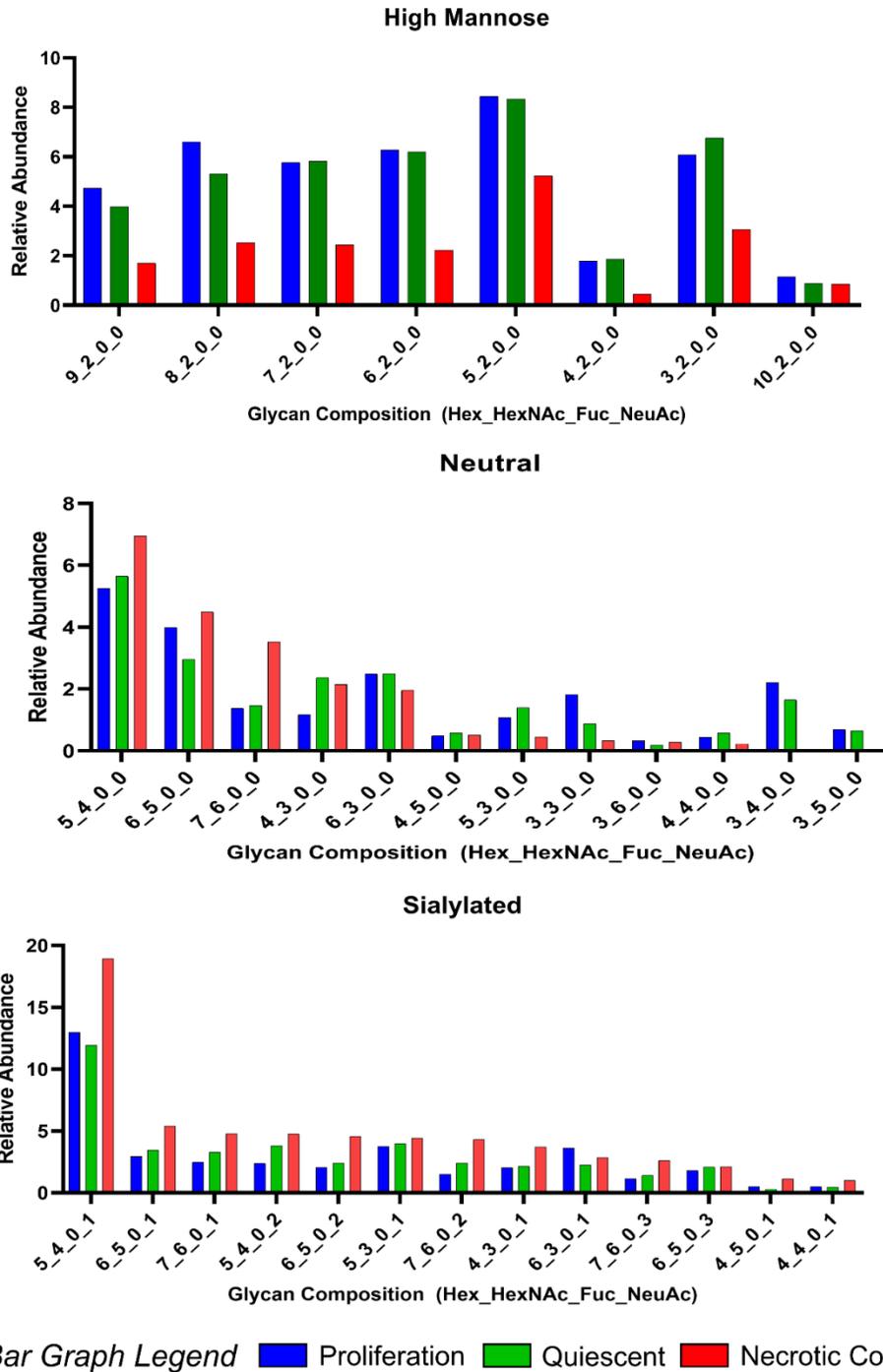
References

1. Siegel, R. L.; Miller, K. D.; Goding Sauer, A.; Fedewa, S. A.; Butterly, L. F.; Anderson, J. C.; Cercek, A.; Smith, R. A.; Jemal, A., Colorectal Cancer Statistics, 2020. *CA: A Cancer Journal for Clinicians* **2020**, *70* (3), 145-164.
2. Holm, M.; Nummela, P.; Heiskanen, A.; Satomaa, T.; Kaprio, T.; Mustonen, H.; Ristimäki, A.; Haglund, C., N-glycomic Profiling of Colorectal Cancer According to Tumor Stage and Location. *PLoS One* **2020**, *15* (6), e0234989.
3. Zhao, Q.; Zhan, T.; Deng, Z.; Li, Q.; Liu, Y.; Yang, S.; Ji, D.; Li, Y., Glycan Analysis of Colorectal Cancer Samples Reveals Stage-Dependent Changes in CEA Glycosylation Patterns. *Clin Proteomics* **2018**, *15*, 9.
4. Holst, S.; Deuss, A. J.; van Pelt, G. W.; van Vliet, S. J.; Garcia-Vallejo, J. J.; Koeleman, C. A.; Deelder, A. M.; Mesker, W. E.; Tollenaar, R. A.; Rombouts, Y.; Wührer, M., N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homebox 1 (CDX1)/Villin mRNA Expression. *Mol Cell Proteomics* **2016**, *15* (1), 124-40.
5. Sethi, M. K.; Fanayan, S., Mass Spectrometry-Based N-Glycomics of Colorectal Cancer. *Int J Mol Sci* **2015**, *16* (12), 29278-304.
6. Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J., Glycosylation in Health and Disease. *Nat Rev Nephrol* **2019**, *15* (6), 346-366.
7. de Leoz, M. L.; Young, L. J.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B., High-mannose Glycans are Elevated During Breast Cancer Progression. *Mol Cell Proteomics* **2011**, *10* (1), M110 002717.
8. Daniotti, J. L.; Vilcaes, A. A.; Torres Demichelis, V.; Ruggiero, F. M.; Rodriguez-Walker, M., Glycosylation of Glycolipids in Cancer: Basis for Development of Novel Therapeutic Approaches. *Front Oncol* **2013**, *3*, 306.
9. Oh, Y. J.; Dent, M. W.; Freels, A. R.; Zhou, Q.; Lebrilla, C. B.; Merchant, M. L.; Matoba, N., Antitumor Activity of a Lectin Targeting Cancer-Associated High-Mannose Glycans. *bioRxiv* **2021**, 2021.04.28.441869.
10. Katt, M. E.; Placone, A. L.; Wong, A. D.; Xu, Z. S.; Searson, P. C., In Vitro Tumor Models: Advantages, Disadvantages, Variables, and Selecting the Right Platform. *Front Bioeng Biotechnol* **2016**, *4*, 12.
11. Vidi, P. A.; Bissell, M. J.; Lelievre, S. A., Three-Dimensional Culture of Human Breast Epithelial Cells: the How and the Why. *Methods Mol Biol* **2013**, *945*, 193-219.
12. Lin, R. Z.; Chang, H. Y., Recent Advances in Three-dimensional Multicellular Spheroid Culture for Biomedical Research. *Biotechnol J* **2008**, *3* (9-10), 1172-84.
13. Caragher, S.; Chalmers, A. J.; Gomez-Roman, N., Glioblastoma's Next Top Model: Novel Culture Systems for Brain Cancer Radiotherapy Research. *Cancers (Basel)* **2019**, *11* (1), 44-91.
14. Weigelt, B.; Ghajar, C. M.; Bissell, M. J., The Need for Complex 3D Culture models to Unravel Novel Pathways and Identify Accurate Biomarkers in Breast Cancer. *Adv Drug Deliv Rev* **2014**, *69-70*, 42-51.
15. Liu, X.; Hummon, A. B., Quantitative Determination of Irinotecan and the Metabolite SN-38 by Nanoflow Liquid Chromatography-Tandem Mass Spectrometry in Different Regions of Multicellular Tumor Spheroids. *J Am Soc Mass Spectrom* **2015**, *26* (4), 577-86.

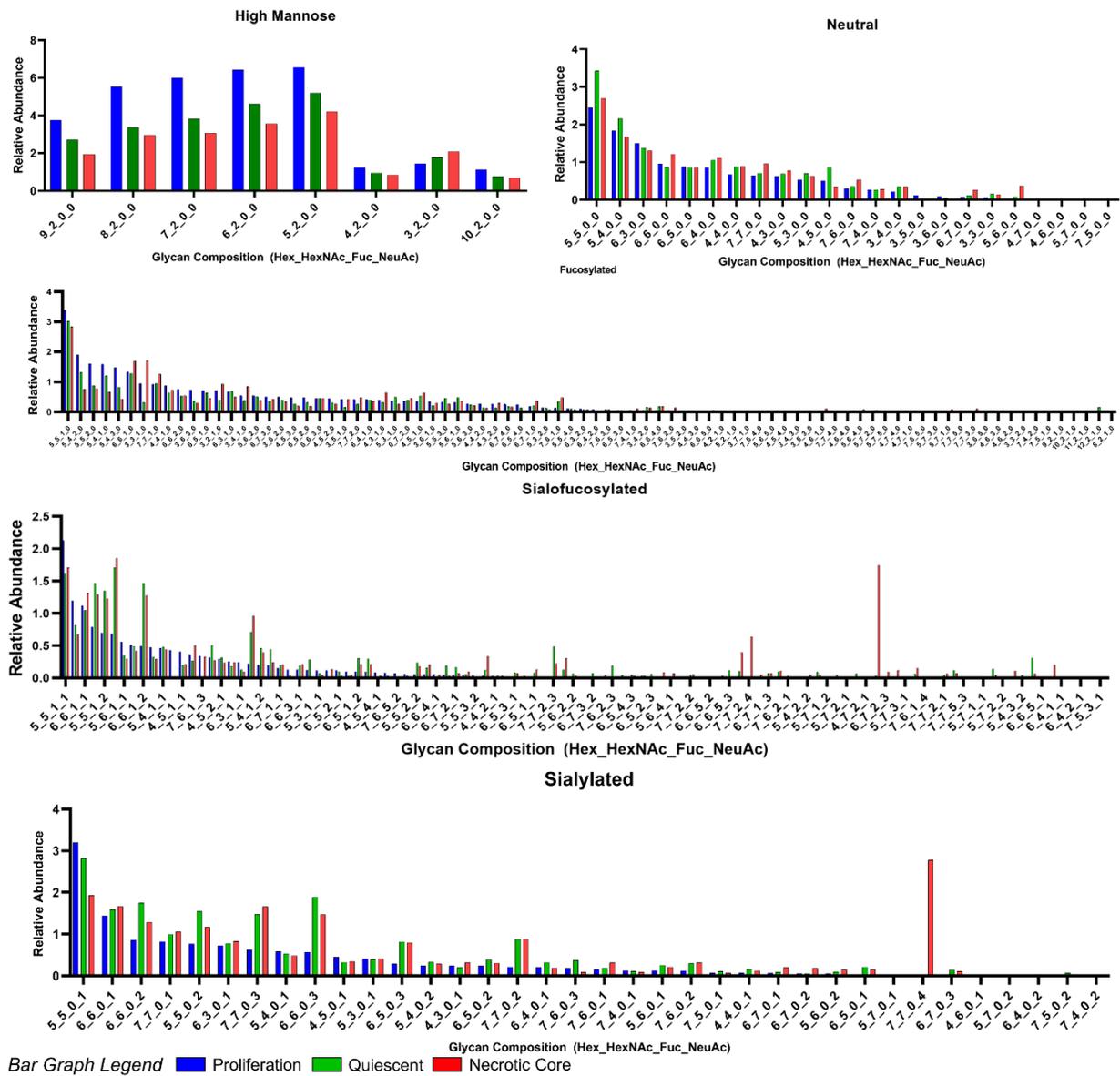
16. Yue, X.; Lukowski, J. K.; Weaver, E. M.; Skube, S. B.; Hummon, A. B., Quantitative Proteomic and Phosphoproteomic Comparison of 2D and 3D Colon Cancer Cell Culture Models. *J Proteome Res* **2016**, *15* (12), 4265-4276.
17. Neef, S. K.; Janssen, N.; Winter, S.; Wallisch, S. K.; Hofmann, U.; Dahlke, M. H.; Schwab, M.; Murdter, T. E.; Haag, M., Metabolic Drug Response Phenotyping in Colorectal Cancer Organoids by LC-QTOF-MS. *Metabolites* **2020**, *10* (12), 494-526.
18. Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatilleke, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem Rev* **2018**, *118* (17), 7886-7930.
19. Kailemia, M. J.; Xu, G.; Wong, M.; Li, Q.; Goonatilleke, E.; Leon, F.; Lebrilla, C. B., Recent Advances in the Mass Spectrometry Methods for Glycomics and Cancer. *Anal Chem* **2018**, *90* (1), 208-224.
20. 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.
21. 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.
22. Li, Q.; Xie, Y.; Xu, G.; Lebrilla, C. B., Identification of Potential Sialic Acid Binding Proteins on Cell Membranes by Proximity Chemical Labeling. *Chem Sci* **2019**, *10* (24), 6199-6209.
23. 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.
24. Wong, M.; Xu, G.; Park, D.; Barboza, M.; Lebrilla, C. B., Author Correction: Intact Glycosphingolipidomic Analysis of the Cell Membrane During Differentiation Yields Extensive Glycan and Lipid Changes. *Scientific Reports* **2020**, *10* (1), 21377.
25. Ozcan, S.; An, H. J.; Vieira, A. C.; Park, G. W.; Kim, J. H.; Mannis, M. J.; Lebrilla, C. B., Characterization of Novel O-Glycans Isolated from Tear and Saliva of Ocular Rosacea Patients. *Journal of Proteome Research* **2013**, *12* (3), 1090-1100.
26. Consortium, T. U., UniProt: the Universal Protein Knowledgebase in 2021. *Nucleic Acids Research* **2020**, *49* (D1), D480-D489.
27. Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.; Bork, P.; Jensen, L. J.; Mering, Christian v., STRING v11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Research* **2018**, *47* (D1), D607-D613.
28. Raudvere, U.; Kolberg, L.; Kuzmin, I.; Arak, T.; Adler, P.; Peterson, H.; Vilo, J., g:Profiler: a Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 update). *Nucleic Acids Research* **2019**, *47* (W1), W191-W198.
29. Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D. A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P.-É.; Li, S.; Xia, J., MetaboAnalyst 5.0: Narrowing the Gap Between Raw Spectra and Functional Insights. *Nucleic Acids Research* **2021**, *49* (W1), W388-W396.
30. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res* **2018**, *46* (W1), W296-w303.

31. Park, S.-J.; Lee, J.; Qi, Y.; Kern, N. R.; Lee, H. S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im, W., CHARMM-GUI Glycan Modeler for Modeling and Simulation of Carbohydrates and Glycoconjugates. *Glycobiology* **2019**, *29* (4), 320-331.
32. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605-1612.
33. Duldulao, M. P.; Lee, W.; Le, M.; Chen, Z.; Li, W.; Wang, J.; Gao, H.; Li, H.; Kim, J.; Garcia-Aguilar, J., Gene Expression Variations in Microsatellite Stable and Unstable Colon Cancer Cells. *J Surg Res* **2012**, *174* (1), 1-6.
34. Yao, K.; Gietema, J. A.; Shida, S.; Selvakumaran, M.; Fonrose, X.; Haas, N. B.; Testa, J.; O'Dwyer, P. J., In Vitro Hypoxia-Conditioned Colon Cancer Cell Lines Derived from HCT116 and HT29 Exhibit Altered Apoptosis Susceptibility and a More Angiogenic Profile In Vivo. *Br J Cancer* **2005**, *93* (12), 1356-63.
35. Taki, T.; Takamatsu, M.; Myoga, A.; Tanaka, K.; Ando, S.; Matsumoto, M., Glycolipids of Metastatic Tissue in Liver from Colon Cancer: Appearance of Sialylated Lex and Lex Lipids1. *The Journal of Biochemistry* **1988**, *103* (6), 998-1003.
36. Park, D. D.; Phoomak, C.; Xu, G.; Olney, L. P.; Tran, K. A.; Park, S. S.; Haigh, N. E.; Luxardi, G.; Lert-Itthiporn, W.; Shimoda, M.; Li, Q.; Matoba, N.; Fierro, F.; Wongkham, S.; Maverakis, E.; Lebrilla, C. B., Metastasis of Cholangiocarcinoma is Promoted by Extended High-Mannose Glycans. *Proc Natl Acad Sci U S A* **2020**, *117* (14), 7633-7644.
37. Brockhausen, I., Mucin-type O-glycans in Human Colon and Breast Cancer: Glycodynamics and Functions. *EMBO Rep* **2006**, *7* (6), 599-604.
38. Yang, J. M.; Byrd, J. C.; Siddiki, B. B.; Chung, Y. S.; Okuno, M.; Sowa, M.; Kim, Y. S.; Matta, K. L.; Brockhausen, I., Alterations of O-Glycan Biosynthesis in Human Colon Cancer Tissues. *Glycobiology* **1994**, *4* (6), 873-84.
39. 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.
40. Chandler, P. D.; Akinkuolie, A. O.; Tobias, D. K.; Lawler, P. R.; Li, C.; Moorthy, M. V.; Wang, L.; Duprez, D. A.; Jacobs, D. R.; Glynn, R. J.; Otvos, J.; Connelly, M. A.; Post, W. S.; Ridker, P. M.; Manson, J. E.; Buring, J. E.; Lee, I. M.; Mora, S., Association of N-Linked Glycoprotein Acetyls and Colorectal Cancer Incidence and Mortality. *PLoS One* **2016**, *11* (11), e0165615.
41. Zhu, W.; Ma, L.; Qian, J.; Xu, J.; Xu, T.; Pang, L.; Zhou, H.; Shu, Y.; Zhou, J., The Molecular Mechanism and Clinical Significance of LDHA in HER2-Mediated Progression of Gastric Cancer. *Am J Transl Res* **2018**, *10* (7), 2055-2067.
42. Hsu, M.-C.; Hung, W.-C., Pyruvate Kinase M2 Fuels Multiple Aspects of Cancer Cells: from Cellular Metabolism, Transcriptional Regulation to Extracellular Signaling. *Molecular Cancer* **2018**, *17* (1), 35.

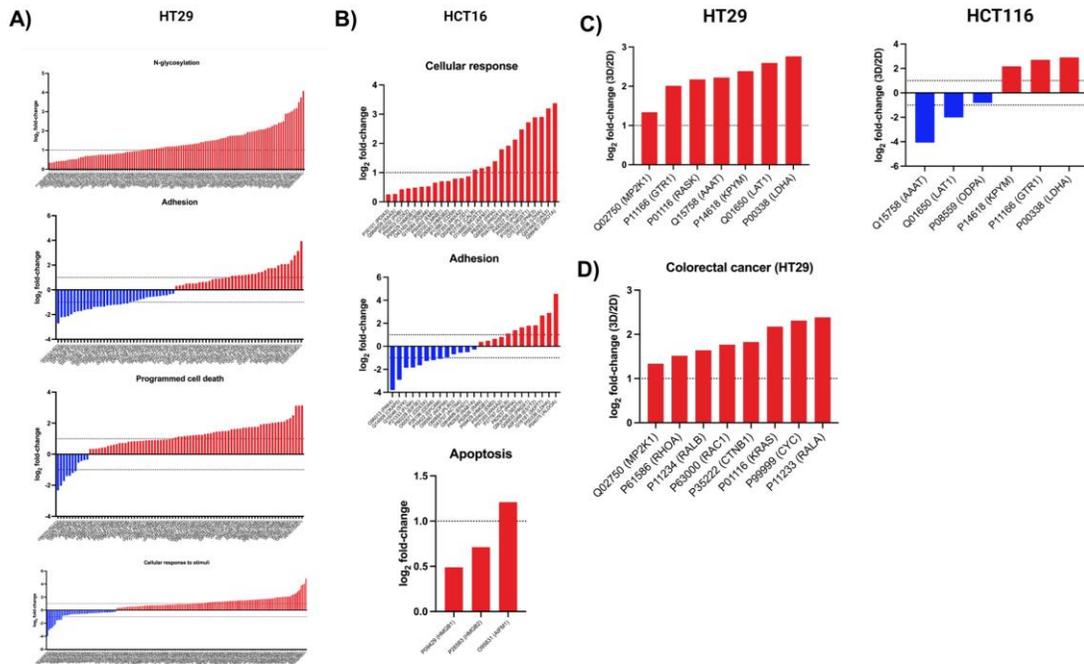
Supplementary Figures



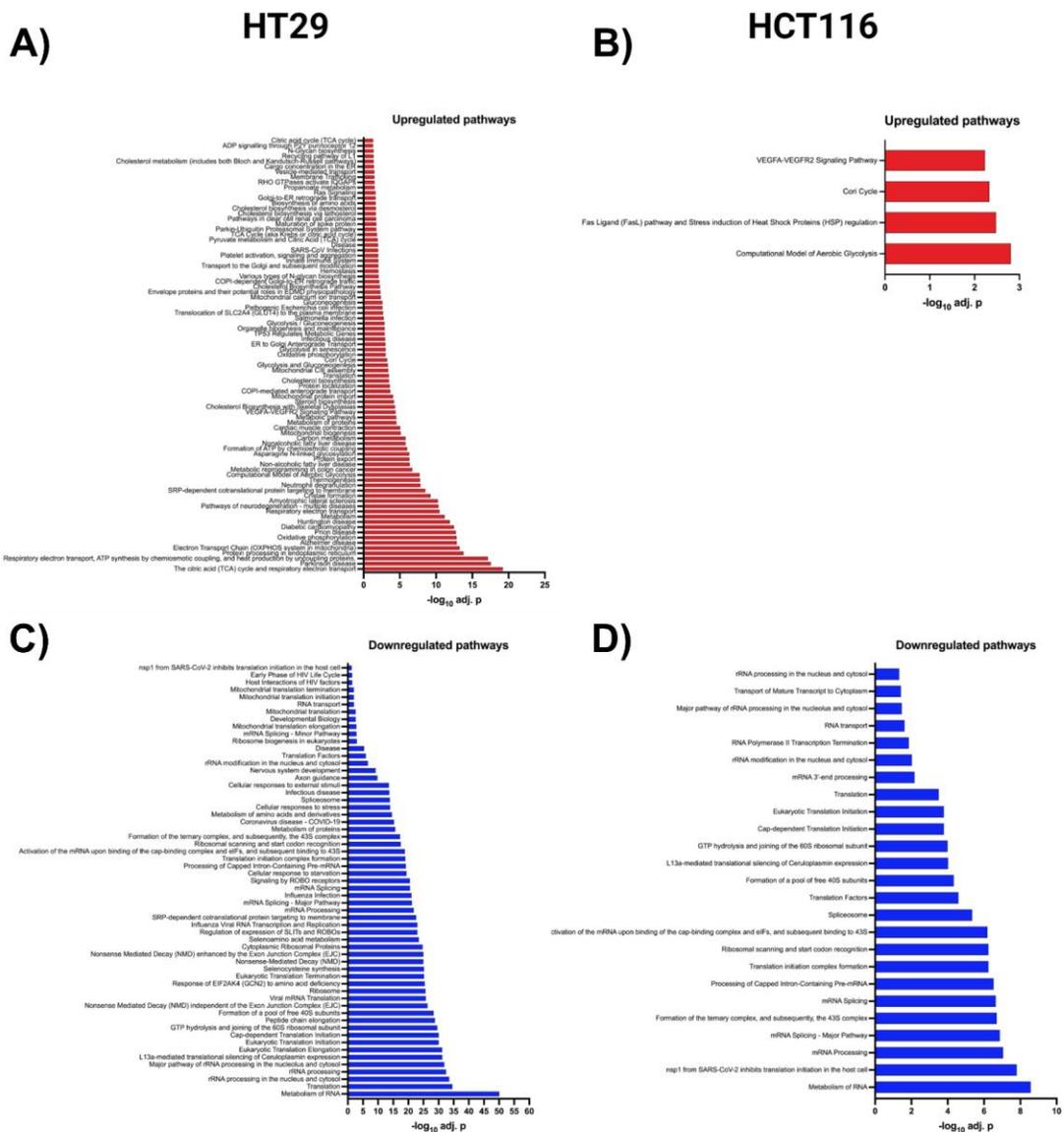
Supplementary Figure 3.1 The relative abundance of N-glycans of each layer where they are grouped into their N-glycan categories for HCT116.



Supplementary Figure 3.2 The relative abundance of N-glycans of each layer where they are grouped into their N-glycan categories for HT29.



Supplementary Figure 3.3 The significantly altered proteomic between the 2D and 3D model where (A,B) Protein expressions of significantly-different proteins involved in N-glycosylation, adhesion, programmed cell death/apoptosis, and cellular responses to stress, in HT29 and HCT116. (C) Protein expressions of significantly-different proteins in 3D- and 2D-cultured (A) HT29 and (B) HCT116 that are implicated in central carbon metabolism in cancer (KEGG id: hsa05230, FDR=0.025 and 0.0139, respectively), and (D) Protein expressions of significantly-different proteins in 3D- and 2D-cultured HT29 that are implicated in colorectal cancer (KEGG id: hsa05210, FDR=0.0268).



Supplementary Figure 3.4. Gene ontology analysis of significantly-different proteins, showing up-regulated (A, B) and down-regulated (C,D) pathways in HT29 and HCT116 3D-cultured cells. In HT29, 340 pathways were up-regulated and 338 pathways were down-regulated. In HCT116, 26 pathways were up-regulated and 113 pathways were down-regulated.

Chapter IV

Functional Characterization of an N-glycosylation Inhibitor Using Integrated Bioinformatics & Mass-spectrometric Glycomic-Driven Glycoproteomic Analysis

Abstract

Protein glycosylation is a complex process that is mediated by multiple glycosyltransferases and glycosidases, resulting in a wide variety of N-glycosylation in cell-surface glycoproteins. Cancer progression has been linked to aberrant glycosylation due to overexpression of several glycosylation enzymes, such as Alpha1-6FucT, GlcNAcT-V, and Alpha2,6-ST I. These enzymes are an underexploited drug target in cancer therapeutics; as such, there is a lack of glycosylation inhibitors with drug-like properties on the market. Thus, we aim to utilize a computer-aided approach in identifying potential glycosylation inhibitors. A network pharmacology approach coupled with *in silico* screening was used to identify a potential inhibitor, pictilisib, from a database of known drugs against several glycosylation-related proteins. A549 cells (non-small cell lung carcinoma) were treated with pictilisib to determine its effect on protein glycosylation. Mass spectrometry-based glycomics assay shows that pictilisib significantly reduces fucosylation and sialylation of N-glycans. Proteomics analysis and *in vitro* assays show significant upregulation of proteins involved in apoptosis and cell-adhesion as well as downregulation of proteins involved in cell cycle regulation, mRNA processing, and protein translation. Site-specific glycoproteomics analysis further shows that glycoproteins with reduced fucosylation and sialylation were involved in apoptosis, cell-adhesion, DNA damage repair, and chemical response processes. To determine how changes in N-glycosylation could affect the dynamics of glycoproteins, we modeled the changes in glycan interactions of the ITGA5-ITGB1 (Integrin alpha 5-Integrin beta-1) complex. We found specific glycosites at the interface of the two proteins that, when fucosylated and sialylated, could form more hydrogen bonds compared to high-mannose types obtained from pictilisib-treated A549 cells. We present for the first time how a drug, pictilisib, affects protein N-glycosylation and the pathways involving these glycoproteins through an integrated multi-OMICS and bioinformatics pipeline.

Introduction

Lung cancer is the leading cause of cancer-related mortalities worldwide ¹. Among males, lung cancer is the leading cause of death in Eastern Europe, Western Asia, Northern Africa, and Asia ². Among females, lung cancer has the highest incidence rates in North America, Northern, Western Europe, and Australia/New Zealand. Cancer incidence and mortality are multiplying worldwide, reflecting several factors: aging, population growth, cancer risk factors, and socioeconomic development. According to the GLOBOCAN 2018 database of 185 countries and 36 cancers, there will be 18.1 million new cases and 9.6 million cancer deaths worldwide ². Out of these, 2,093,876 cases (11.6%) and 1,761,007 (18.4%) deaths for both sexes are due to lung cancer. Additionally, lung cancer is the leading cause of cancer death among men in 93 countries and women in 28 countries.

Protein glycosylation is one of the most complex and most frequent post-translational modifications and is involved in many cellular interactions such as host-pathogen interactions, cell differentiation and trafficking, and intra- and intercellular signaling ³. Protein glycosylation is a complex process that starts at the endoplasmic reticulum and is further processed in the Golgi apparatus. In the Golgi apparatus, the glycans are further processed to achieve the diversity and complexity of final glycan structures through a series of steps involving glycosyltransferases and glycosidases. Overexpression of these glycan-processing enzymes is usually observed in cancer cells, resulting in enhanced expression of related glycan structures. For example, the enzymes Alpha1-6FucT, B4GALT2, MAN1A2, and MAN2A1 are overexpressed in lung cancer tissue samples ⁴. Likewise, glycans corresponding to these enzymes are also overexpressed in lung cancer tissues ⁵. Additionally, aberrant glycosylation also leads to increased biosynthesis of various tumor antigens, such as Sialyl Lewis X, which serves as a ligand for the cell-adhesion molecule selectin.

This antigen is also involved in the adhesion of cancer cells to vascular endothelium and hematogenous metastasis. Furthermore, increased glycosylation is also implicated in the loss of E-cadherin, which is essential in the metastatic dissemination of cells⁶. Thus, cancer progression is also associated with changes in the glycosylation of cell surface proteins involved in the loss of cell to cell-adhesion and increased metastatic potential. Furthermore, altered glycosylation is also correlated with the other hallmarks of cancer such as enhanced proliferation, angiogenesis potential, replicative immortality, metastatic potential, apoptosis and tumor suppression⁷.

Several glycosyltransferases have been associated as cancer biomarkers³. A glycosyltransferase used as a biomarker is UDP-N-acetyl-D-glucosamine: N-acetylglucosamine transferase V (GlcNAcT-V), which catalyzes β 1-6 branching of N-glycans. Increased β 1-6 branching, due to GlcNAcT-V overexpression, has been observed in breast carcinoma⁸. Sialyltransferases are glycosyltransferases that are abnormally expressed in cancers and are implicated in carcinogenesis, progression, and metastasis⁹⁻¹¹. Overexpression of α 2-3 sialyltransferase III (ST3Gal-III) in pancreatic cancer has been implicated in pancreatic tumor progression. Overexpression of α 2-6 sialyltransferase I (ST6GalNAc-I) was related to poor patient survival in colorectal carcinoma patients¹². As such, glycosyltransferases and glycosidases are underexploited drug targets for cancer therapeutics and there is a relative lack of small molecule inhibitors of these enzymes with drug-like properties. Esko and Bertozzi (2009) classified glycosylation inhibitors and their targets into seven classes: metabolic inhibitors, which target the formation of nucleotide sugars, tunicamycin which target dolichol-PP-GlcNAc formation (biosynthesis of N-glycans), plant alkaloids that inhibit processing of glycosidases, substrate analogs which are specific towards glycosyltransferases and glycosidases, glycoside primers which divert assembly of glycans from endogenous acceptors towards exogenous primers, and

tagged monosaccharides which target several different biosynthesis pathways¹³. Examples of plant alkaloids that inhibit glycosidases are tunicamycin and glucosamine. Both of these compounds induce inhibition of protein N-glycosylation by blocking the GlcNAc phosphotransferase-catalyzed transfer of N-acetylglucosamine-1-phosphate from UDP-GlcNAc to dolichol-P, which results in decreased production of dolichol-PP-GlcNAc. In combination with anticancer drugs, tunicamycin has also been shown to be cytotoxic against multidrug-resistant human ovarian cystadenocarcinoma cells by inhibiting protein and glycoprotein syntheses¹⁴. The flavonoid Scutellarein is also shown to inhibit the proliferation of the non-small cell lung carcinoma (NSCLC) cell line A549 through inhibition of ERK and NFkB via the EGFR pathway¹⁵.

In this study, we utilized an integrated computational approach – network pharmacology and *in silico* docking – to identify potent glycosylation inhibitors. Using this approach, we were able to identify pictilisib as a potent glycosylation inhibitor in A549 cells. Further analysis using mass-spectrometric glycomics, proteomics, and glycoproteomics shows how affecting protein N-glycosylation could affect cancer pathways.

Materials and Methods

Network Pharmacology

A ligand database was prepared by downloading the structure data files (.sdf) from several online databases, such as the Comparative Toxicogenomics database (<http://ctdbase.org/>)¹⁶, STITCH database (<http://stitch.embl.de/>)¹⁷, GeneCards (<https://www.genecards.org/>)¹⁸, Drug Gene Interaction database (<http://www.dgidb.org/>)¹⁹, and Protein Databank (<https://www.rcsb.org/>)²⁰, and 185 compounds predicted to bind or interact with glycosylation enzymes from DrugBank (<https://go.drugbank.com/>)²¹. Gene-drug interactions were predicted using the STITCH database and Comparative Toxicogenomics database and then visualized using

Cytoscape²².

In Silico Docking

All compounds from the ligand database were loaded onto PyRx²³ and minimized using the Universal Force Field²⁴ as implemented in Open Babel²⁵. The enzymes GlcNAcT-V, Alpha2,6-ST I, and Alpha1-6FucT were selected as the drug target for this study due to the availability of their 3D crystal structures in PDB. The enzyme GlcNAcT-V (PDB ID: 5ZIC, 2.10 Å)²⁶ was downloaded as a complex with its acceptor sugar ,2-acetamido-2-deoxy-beta-D-glucopyranose-(1-2)-6-thio-alpha-D-mannopyranose-(1-6)-beta-d-mannopyranose. Alpha2,6-ST I (PDB ID: 4JS2, 2.30 Å)²⁷ was downloaded as a complex with cytidine monophosphate. Human Alpha1-6FucT (PDB ID: 2de0)²⁸ was homology modelled from *Caenorhabditis elegans* POFUT1 (PDB ID: 3ZY6, 1.91 Å)²⁹ in complex with GDP-fucose, using SWISS-MODELLER³⁰. These protein structures were prepared for docking by using the Dockprep protocol in Chimera³¹. The prepared protein structures were loaded in PyRx as macromolecule receptors.

In silico screening methods were performed in PyRx²³ using the AutoDock VINA docking protocol³² at exhaustiveness level 8. Validation of the docking protocol was done by redocking the ligands complexed with their respective enzymes using precise grid box parameters. After docking validation, all compounds in the ligand database were screened against each of the three enzymes. The compounds were ranked according to VINA-predicted binding energy (kcal/mol). The top binding molecules against each enzyme were visualized for residue interactions with the target enzyme using Discovery StudioTM (Dassault Systemes). From the screening, pictilisib was found to have high compound cross-reactivity (binding to multiple enzyme targets) and a high number of network interactions and was selected for further *in vitro* studies.

Cell Culture

The cell line A549 (CCL-185TM) was obtained from the American Type Culture Collection (ATCC). A549 cells were grown in 20 mL cultured in RPMI 1640 medium (Gibco) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Thermo Scientific) in T75 flasks. The media was changed every other day. All cells were grown in at least three biological replicates and maintained in a humidified incubator at 37 °C and in an atmosphere of 5% CO₂.

Dose-Response Assay

A549 cells were seeded into 96-well plates at 3000 cells/well. The plates were incubated at 37 °C, 5% CO₂, for 24 hours to allow attachment and proliferation. After which, the cells were treated with half-log dilutions of pictilisib (SelleckChem) and negative control (1% v/v DMSO). The cells were incubated with the compounds at 37 °C, 5% CO₂, for 24 hours. Cell viability was detected using CellTiter⁹⁶ aqueous MTS assay reagent (Promega) following manufacturer's instructions. The IC₅₀ value was calculated using GraphPad Prism using % viability as input values for each log [pictilisib] concentration. Assays were done in triplicate.

Cell Cycle Assay

A549 cells were seeded into 100 mm plates. Upon reaching approximately 80% confluency, the cells were treated with 4 μM pictilisib (final concentration), 0.1% v/v DMSO (negative control) or 100 μM docetaxel (positive control) for 24 hours at 37 °C, 5% CO₂. Cell cycle assay was performed using CellometerTM PI Cell cycle kit (Nexcelom) according to manufacturer instructions. Data was acquired using Cellometer Vision CBA 5 (Nexcelom) using the protocol CBA_Cell Cycle-PI660 nm, with an exposure time of 15000 msec. The acquired image cytometry data were analyzed using FCS Express 7.0. Cell gating was adjusted based on negative control. Assays were done in triplicate.

Apoptosis Assay

A549 cells were seeded into 100 mm plates. Upon reaching approximately 80% confluency, the cells were treated with 4 μ M pictilisib (final concentration), 0.1% v/v DMSO (negative control) or 100 μ M docetaxel (positive control) for 24 hours at 37 °C, 5% CO₂. Cell cycle assay was performed using Annexin V-FITC/PITM Apoptosis kit (Nexcelom) according to manufacturer instructions. Data was acquired using Cellometer Vision CBA 5 (Nexcelom), using the protocol CBA_Annexin V+PI assay, with an F1 exposure time of 8000 msec and F2 exposure time of 20000 msec. Acquired data were analyzed using FCS Express 7.0. Cell gating was adjusted based on negative and positive controls. Assays were done in triplicate.

Scratch Assay

A549 cells were seeded into 6-well plates and allowed to grow to confluency at 37 °C, 5% CO₂. Cell surface scratches were made using P200 pipette tips, then washed twice with PBS to remove the debris. The plates were supplemented with RPMI media (2% FBS, 1% penicillin-streptomycin) to reduce the effects of cell proliferation. Cells were treated with a final concentration of 4 μ M pictilisib or with 0.1% v/v DMSO (negative control) for 48 hours. Micrographs were taken starting from 0 hr and every 12 hours thereafter. Wound size areas were measured using ImageJ software³³ and reported relative to initial wound size. Assays were done in triplicate.

Trans-Well Migration Assay

A549 cells were grown in 100 mm plates until reaching approximately 80% confluency. Then, the cells were treated to a final concentration of 4 μ M pictilisib or with 0.1% v/v DMSO (negative control) for 24 hours. After which, the cells were harvested using trypsin and adjusted to 50,000 cells per mL in complete media. One (1) mL of the resulting suspension was pipetted

into the top compartment of a trans-well plate. The bottom chamber was filled with RPMI media (1% penicillin-streptomycin) without FBS to establish chemotaxis. The plates were subsequently incubated for 3 hours at 37 °C, 5% CO₂, to allow cell migration. The plates were washed twice with HBSS, fixed with 1% formaldehyde for 5 minutes, and the bottom compartment stained with crystal violet. Cells that migrated through the trans-well membrane were visualized in micrographs and manually counted using ImageJ³³ analysis software.

Cell Treatment and Glycan, Protein, and Glycoprotein Enrichment

Cells were grown in T75 flasks until reaching approximately 80% confluency. The cells were treated to a final concentration of 4 μM pictilisib or with 0.1% v/v DMSO (negative control) for 24 hours. For the glycomic and proteomic mass-spectrometric assay, cells were grown in triplicate T75 flasks for each group. For the glycoproteomic mass-spectrometric assay, cells were grown in 15 replicate T75 flasks for each group. After culturing, the general protocol for all mass-spectrometric analyses was taken from Li *et al.*, 2020³⁴.

Glycomics Assay

N-Glycan profiling was performed using an Agilent 6200 series nanoHPLC-Chip-QTOF-MS (Agilent) with an Agilent 6210 time-of-flight mass spectrometer. The chip (glycan chip II, Agilent) contained a 9 mm × 0.075 mm i.d. enrichment column coupled to a 43 mm × 0.075 mm i.d. analytical column; both are packed with 5-μm porous graphitized carbon (PGC). N-glycan samples were reconstituted in 40 μL of water and 5 μL of the resulting solution was used for injection into the LC-MS/MS system. Upon injection, the sample was loaded onto the enrichment column using 3% ACN containing 0.1% formic acid (FA, Fluka, St. Louis, MO). After the analytical column was switched in-line, the nanopump delivered a gradient of 3% ACN with 0.1% FA (solvent A) and 90% ACN with 1% FA (solvent B). The sample was delivered by the capillary

pump to the enrichment column at a flow rate of 3 $\mu\text{L}/\text{min}$ and separated on the analytical column by the nanopump at a flow rate of 0.3 $\mu\text{L}/\text{min}$ using a gradient optimized for N-glycans (0% B, 0–2.5 min; 0 to 16% B, 2.5–20 min; 16 to 44% B, 20–30 min; 44 to 100% B, 30–35 min; and 100% B, 35–45 min) followed by 20-min equilibration for pure A. Tandem MS spectra were acquired via collision-induced dissociation (CID).

Analysis of the N-glycan data was performed using MassHunter Qualitative Analysis Software B.07.00 (Agilent Technologies). Matching of the monoisotopic masses obtained was done against our in-house database for glycan composition identification and subsequently verified through their corresponding MS/MS spectra. The relative abundance of each glycan in a sample was determined using the peak area of all glycans from extracted ion chromatograms. Comparison between relative abundances of primary N-glycan types - high-mannose, undecorated, fucosylated, sialylated, and sialofucosylated glycans - was done by adding the relative abundances of each glycoform belonging to a specific glycan type. Further comparison of each glycoform was done using multiple t-tests (GraphPad Prism 8) at a significance level of $\alpha=0.05$. Significantly-different N-glycans were mapped on the N-glycan biosynthesis pathway based on the known biosynthetic sequence³⁵.

Proteomics and Glycoproteomics Assay

The pellets containing membrane proteins were reconstituted with 60 μL of 8 M urea and sonicated for 20 minutes for denaturation. Two microliters (2 μL) Dithiothreitol (DTT, 550 mM in 50 mM NH_4HCO_3) was then added to the samples and incubated for 50 minutes at 55 $^\circ\text{C}$. The free cystine was alkylated with 4 μL of iodoacetamide (450 mM) for 20 minutes in the dark at ambient temperature. To quench the reaction, 420 buffer (50 mM NH_4HCO_3) was added. Trypsin (10 μL , 0.1 mg/mL) was then added to the mixture and tryptic digestion was performed at 37 $^\circ\text{C}$

for 18 hours. The resulting peptides were purified using a C-18 solid-phase extraction cartridge and dried before LC-MS/MS analysis. To enrich for glycopeptides, the tryptic digests were cleaned up using HILIC solid-phase extraction and dried before LC-MS/MS analysis. The purified peptides were adjusted to 0.5 $\mu\text{g}/\mu\text{L}$ while the glycopeptides were adjusted to 0.2 $\mu\text{g}/\mu\text{L}$ before injection using Pierce BCA assay kit following manufacturer's instructions (ThermoFisher).

The proteomics and glycoproteomics samples were characterized using an UltiMate™ WPS-3000RS nanoLC system coupled with an Orbitrap Fusion Lumos MS system (ThermoFisher Scientific). One (1) μL of each sample was injected, and the analytes were separated using an Acclaim™ PepMap™ 100C18 LC Column (3 mm, 0.075 mm x 250 mm, ThermoFisher Scientific) at a flow rate of 300 nL/min. Water containing 0.1% formic acid and 80% acetonitrile containing 0.1% formic acid were used as solvents A and B, respectively. MS spectra were collected with a mass range of m/z 600–2000 at a rate of 1.5 s per spectrum in positive ionization mode. The filtered precursor ions in each MS spectrum were subjected to fragmentation through 30% higher-energy C-trap dissociation (HCD) using nitrogen gas as carrier.

The mass spectrometry data were analyzed using Byos workflow (Protein Metrics). For qualitative analysis in Byonic (Protein Metrics), proteins were identified against the human proteome database³⁶ using a precursor mass tolerance of 20 ppm and fragment mass tolerance of 10 ppm. The digestion parameters used included C-terminal cleavage by trypsin (K and R cleavage sites) with at most two missed cleavages. The following peptide modifications were included: carbamidomethyl @ C, oxidation @ M, deamidation @ N and Q, acetylation at protein N-terminal, Gln to pyro-Glu at N-terminal Q, Glu to pyro-Glu at N-terminal E. Protein IDs were filtered at 1% FDR. To identify the glycoproteins and glycoforms, an additional search was performed in Byonic using an in-house N-glycan database. Quantification for each protein was done in Byologic

(Protein Metrics) by quantifying the XIC area sum of the top 3 most abundant peptides. XICs were then normalized to sum total before statistical analysis. On the other hand, glycoform quantification was normalized to each protein's glycosite to yield the percentage occupancy of a particular glycoform.

Gene Ontology Analysis

To identify significantly-different proteins and glycopeptides, multiple t-tests were conducted in GraphPad Prism using an FDR approach (FDR=5%). Significantly over- and under-expressed protein IDs were annotated using Gene Set Enrichment Analysis ³⁷. Similarly, the glycopeptides were annotated using g:Profiler ³⁸ to yield significantly enriched pathways and then plotted as a heatmap in GraphPad Prism.

Glycoprotein Molecular Modeling and Molecular Dynamics

The top glycoproteins were modeled to visualize the effects of changes in glycosylation to protein dynamics and interactions. Interesting glycoproteins were modeled for visualization of the changes in glycosylation to protein dynamics and interactions. Specifically, the proteins ITA5 and ITB1 were selected. Glycoforms in each glycosite were selected based on the highest fold-change between the pictilisib- and negative control-treated cells from the glycoproteomics results. The crystal structures of selected proteins were downloaded from PDB, and then glycans were attached to these proteins using CHARMM-GUI Glycan modeller ³⁹. The system was solvated using the TIP3P model, and 150 mM KCl was added. The CHARMM36 force field was used for both proteins and carbohydrates ⁴⁰. The resulting molecular dynamics input files were used to simulate glycoprotein dynamics for one ns (10000 fs/timestep) in DOST-ASTI High-Performance Computing (HPC) Cluster, Quezon City, Philippines. Simulations were visualized, and the number of interacting hydrogen bonds between glycans and proteins was analyzed using VMD ⁴¹.

Results

Pictilisib was Predicted to Interact with and Inhibit Glycosylation Enzymes Using In Silico Docking and Network Pharmacology

A combined network pharmacology and *in silico* docking approaches were used to identify potential interactors of protein N-glycosylation-related proteins. Several gene-drug interaction databases were surveyed - DrugBank (<https://go.drugbank.com/>), Comparative Toxicogenomics database (<http://ctdbase.org/>), STITCH database (<http://stitch.embl.de/>), GeneCards (<https://www.genecards.org/>), Drug Gene Interaction database (<http://www.dgidb.org/>), and Protein Databank (<https://www.rcsb.org/>) - resulting in 185 predicted glycosylation interactors, that were mapped against 356 glycosylation-related proteins and enzymes (**Figure 4.1A** and **Supplementary Table 4.1**). From this set of compounds, pictilisib was selected due to its high degree of interactions (**Figure 4.1B**). Specifically, pictilisib was predicted to lower the expression of the glycosyltransferase genes B3GALNT1 and B4GALT2 and glycosidase MAN1A1 through interactions with PIK3CA ¹⁶.

Additionally, the compounds were screened and docked onto available crystal structures of three glycosylation proteins - Alpha1-6FucT, Alpha2,6-ST I, and GlcNAcT-V - to predict binding affinities. Here, pictilisib was predicted to bind to the active sites of Alpha1-6FucT directly, Alpha2,6-ST I, and GlcNAcT-V with higher binding affinity than the natural substrate (**Figure 4.1C** and **Supplementary Table 2**). Residue interaction analysis suggests potential pictilisib interactions with critical residues in each enzyme's active site. Against Alpha1-6FucT (Alpha1-6FucT), pictilisib formed Pi-cation interactions with Arg365, hydrogen-bonding interactions with His363, and Pi-Pi T-shaped interactions with Tyr220 (**Figure 4.1D**). In a similar docking experiment by Manabe et al. ⁴², the diphosphate group of GDP-fucose was predicted to form hydrogen bonds with Gly221, Arg365, Ser469, and Gln470. Additionally, the residues His363

sidechain and Tyr250 backbone were shown to tether the guanosine moiety with hydrogen bonds⁴³. Pictilisib also formed hydrogen bonding interactions with Gln235, Pi-Pi T-shaped interactions with His370, and van der Waals interactions with Ala363 (**Figure 4.1E**). Due to these predicted interactions, pictilisib was chosen to modulate glycosylation for further *in vitro* studies. A plausible reaction mechanism of Alpha2,6-ST I suggested by Kuhn et al. shows that His370 could act as the catalytic base for the deprotonation of the 6'-hydroxyl group of the acceptor N-glycan, leading to S_n2 attack of the C2 atom of Neu5Ac²⁷. On the other hand, pictilisib formed hydrogen bonds with Trp401, Asp378, and Leu372 of GlcNAcT-V (**Figure 4.1F**). The sulfur atom in pictilisib also formed Pi-sulfur interactions with Phe380 and Lys554. Nagae et al. (2018) found that these two aromatic residues, Phe380 and Trp401, were found to make contact with the acceptor sugar with Trp401 restraining the conformation of the α1,6-branch²⁶. Most of the top ligands also formed hydrogen bonds with Lys554. In the crystal structure, this residue interacts with the acceptor sugar. These residues - Phe390, Trp401, and Lys554 - are also found in the acceptor substrate binding site for MGAT-IX, suggesting that these residues are relevant in acceptor sugar recognition.

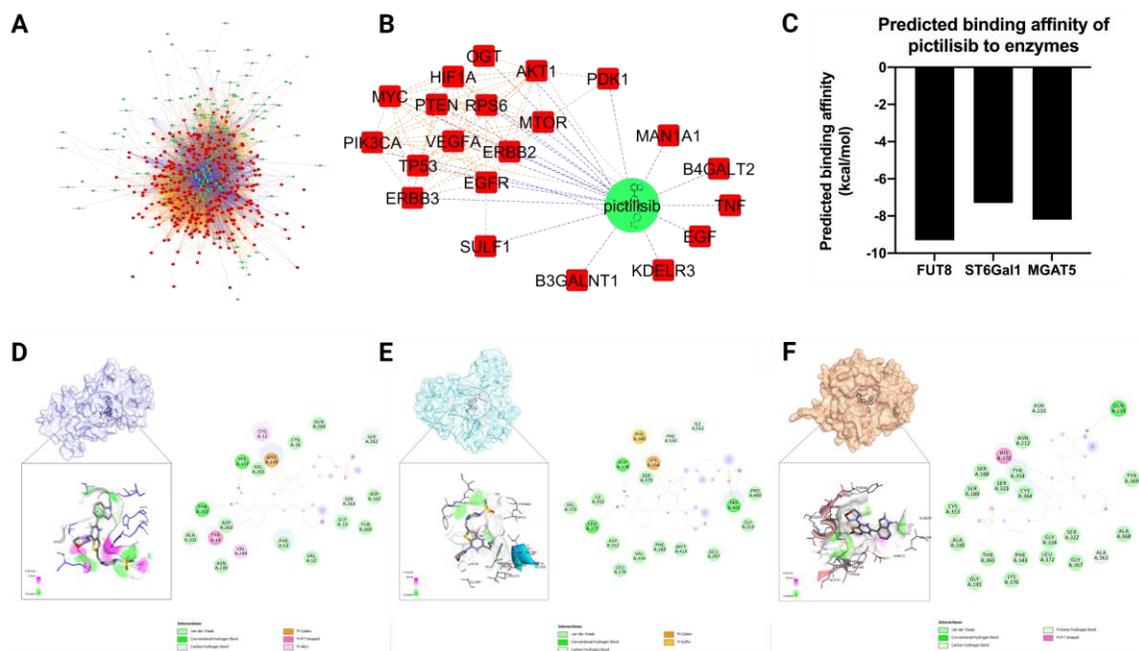


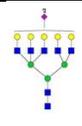
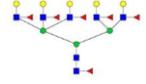
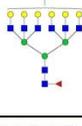
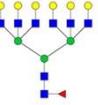
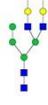
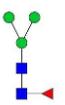
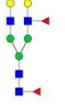
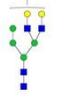
Figure 4.1. Pictilisib was predicted to interact and inhibit several glycosylation-related genes through network pharmacology and *in silico* binding approach. (A) drug-gene interaction network of selected glycosylation targeting compounds. (B) Sub-network of pictilisib drug-gene interactions. (C) The binding affinity of pictilisib against Alpha1-6FucT, Alpha2,6-ST I, and GlcNAcT-V. (D-F) Docking conformation and residue interactions of pictilisib with Alpha1-6FucT, GlcNAcT-V, and Alpha2,6-ST I, respectively.

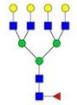
Pictilisib was Validated to Reduce the Relative Abundance of Fucosylated and Sialylated N-Glycans

To determine the effect of changes in protein glycosylation brought by pictilisib treatment, *in vitro* assays in A549 were performed. Prior, dose-response cytotoxic assay and preliminary drug-titration assay after 24 hours of treatment were conducted to determine nontoxic drug concentrations that could still affect protein glycosylation (**Supplementary Figures 4.1 and 4.2**). Upon optimizing assay conditions, A549 cells were treated with pictilisib (4 μ M) for 24 hours then subjected to glycoalyx profiling using the previously published mass-spectrometric method³⁴. Glycomics profiling with mass spectrometry allows for comprehensive and reproducible analysis of the glycan composition of the cell's glycoalyx, with treatment with pictilisib or vehicle control (**Figure 4.2A, Supplementary Table 4.3, and Supplementary Figure 4.3**). Comparing the sum of the relative abundances of the primary N-glycan types - high-mannose, undecorated, fucosylated, sialylated, and sialofucosylated - shows that pictilisib treatment significantly reduced the total relative abundances of fucosylated and sialylated N-glycans (**Figure 4.2B**). A total of 138 glycans were quantified, of which 36 were found to be significant ($p < 0.05$). Closer inspection of these N-glycans shows that total fucosylated complex- and high-mannose-type N-glycans were significantly reduced in pictilisib treatment, while both total sialylated complex- and hybrid-type N-glycans were significantly reduced (**Figure 4.2B**). Specifically, the N-glycan compositions

Hex₆HexNAc₄NeuAc₁, Hex₆HexNAc₄NeuAc₂, Hex₃HexNAc₂Fuc₁, Hex₅HexNAc₄Fuc₂, Hex₇HexNAc₆Fuc₁, Hex₈HexNAc₇Fuc₆, Hex₈HexNAc₇Sia₂, Hex₉HexNAc₈Fuc₁, and Hex₉HexNAc₈Fuc₁NeuAc₂ were found to be very significantly under-expressed in pictilisib-treated cells (**Table 4.1**).

Table 4.1. Glycan composition, putative structures, and log 2-fold-change (and q-values) of highly significantly-different glycans in pictilisib-treated A549 cells compared to vehicle control.

Glycan composition	Putative structure	log ₂ fold-change	-log ₁₀ q-value
Hex ₈ HexNAc ₇ NeuAc ₂		-2.4214	1.6249
Hex ₈ HexNAc ₇ Fuc ₆		-1.6756	1.6249
Hex ₉ HexNAc ₈ Fuc ₁ NeuAc ₂		-1.3536	1.5544
Hex ₉ HexNAc ₈ Fuc ₁		-1.3269	1.5403
Hex ₆ HexNAc ₄ NeuAc ₂		-1.0587	1.6249
Hex ₃ HexNAc ₂ Fuc ₁		-1.0000	1.4854
Hex ₅ HexNAc ₄ Fuc ₂		-0.7675	1.6249
Hex ₆ HexNAc ₄ NeuAc ₁		-0.5385	1.6249

Hex7HexNAc6Fuc1		-0.3078	1.6249
-----------------	---	---------	--------

Mapping these N-glycan compositions to the known N-glycan biosynthetic pathway shows potential glycosylation enzyme reactions that could be inhibited due to pictilisib-treatment (**Figure 2.2C**), specifically those glycosylation reactions involving the addition of fucose and sialic acid residues. These significantly underexpressed N-glycans represent several known cancer-related N-glycan epitopes, such as Lewis and Sialyl Lewis antigens, core fucosylation, and α 2,6-sialylated lactosamine ⁴⁴.

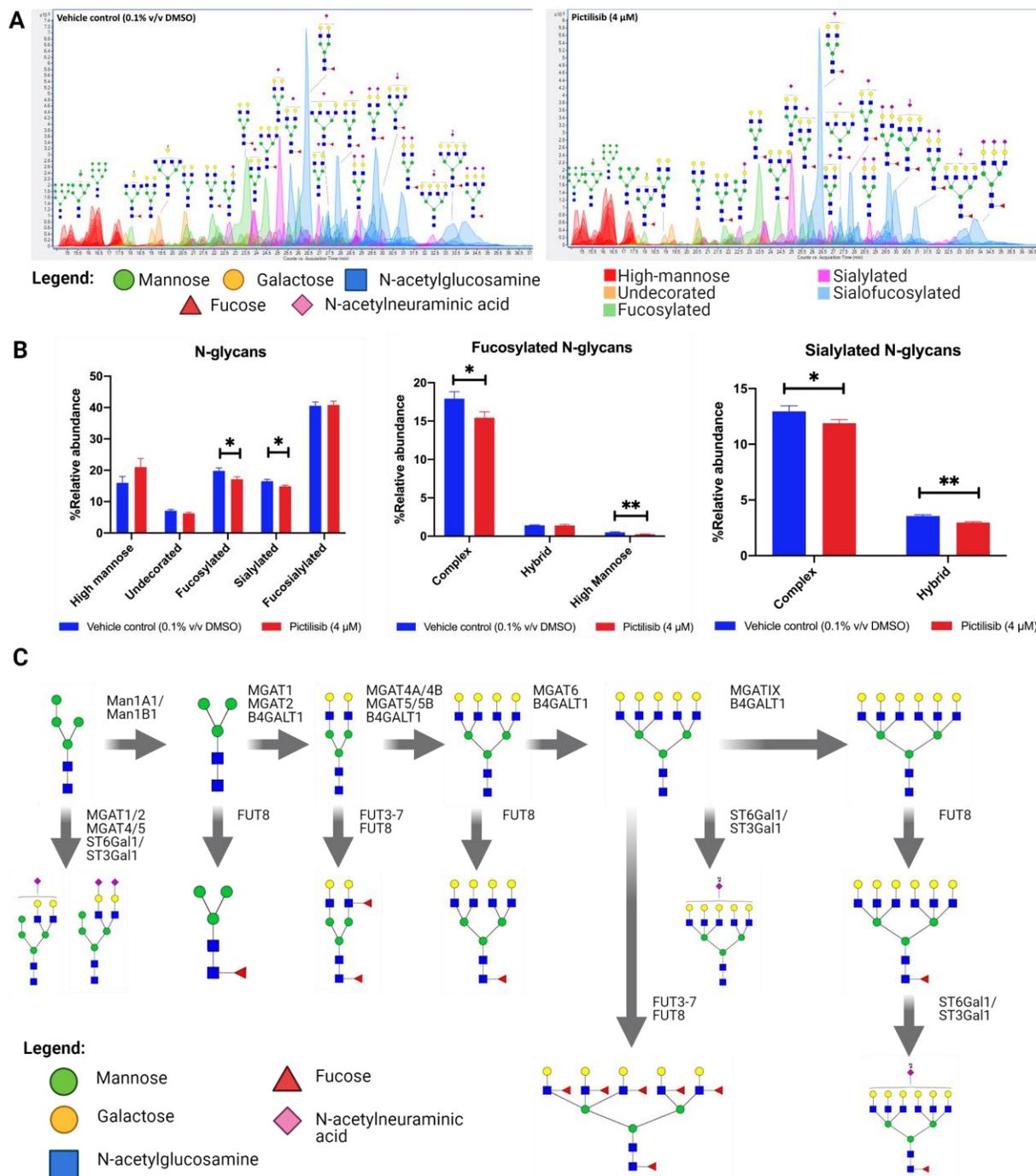


Figure 4.2. Pictilisib significantly reduced the relative abundance of fucosylated and sialylated N-glycans. (A) Glycan-annotated extracted ion chromatograms (EICs) of N-glycomes of vehicle control and pictilisib-treated A549 cells. (B) Relative abundances of N-glycan types in vehicle control- and pictilisib-treated A549 cells. (C) Biosynthetic map showing the abundance of each

significantly different N-glycan (Table 1), N-glycan precursor, and known enzymes catalyzing the glycosylation reaction.

Proteomic Analysis Shows Up-regulated Pathways Involving Apoptosis and Cell-Adhesion, and Down-Regulated Pathways Involving Cell Cycle Process, mRNA Processing, and Protein Translation

To validate the bioactivity effects of pictilisib on A549, we conducted *in vitro* assays coupled with label-free quantitative proteomics to identify specific pathways targeted by pictilisib. Protein was filtered by having Byologic score higher than or equal to 100 and having two unique peptides per protein. Protein intensities were reported as the sum of the top 2 peptides for each protein, normalized to the total intensity per sample. The dataset was further filtered based on the presence of specific protein in at least two replicates per group and then analyzed using multiple t-tests ($\alpha=0.05$) (**Supplementary Figure 4.4**). Based on the proteomics, 1518 proteins were quantified (**Supplementary Figure 4.4**), and 380 proteins were significantly different (p -value <0.05 , **Figure 4.3A**). Gene set enrichment analysis of these significantly different proteins showed interesting biological processes affected by pictilisib treatment, such as upregulation of apoptosis and biological adhesion processes and downregulation of cell cycle processes (**Figures 4.3B, 4.3C, and Supplementary Figure 4.5 and 4.6**). *In vitro* apoptosis and cell cycle assays verify that indeed pictilisib-treatment induced apoptosis (**Figure 4.3D and Supplementary Figure 4.7**) and G0/G1 cell cycle arrest (**Figure 4.3E**) in A549 cells. Correspondingly, quantification of specific apoptosis, cell cycle, and DNA damage-related proteins show significant differences in key proteins involved in these pathways (**Figures 4.3F, 4.3G, and 4.3H**). Likewise, the effect of pictilisib on cell migration was verified using both scratch assay and trans-well migration assay, with pictilisib causing a significant reduction in cell migration (**Figures 4.3I and 4.3J, Supplementary Figures 4.9 and 4.10**). Proteomics analysis shows that the mechanism affecting

Figure 4.3. Pictilisib treatment significantly affected pathways involving ECM interactions and migration, and cell death and proliferation, in A549. (A) Volcano plot of differentially expressed proteins in pictilisib-treated A549 cells. (B) Gene-set enrichment analysis of pre-ranked protein expression profiles of pictilisib- vs. vehicle control-treated cells. (C) Processes involved in apoptosis regulation and biological adhesion were up-regulated, while processes involved in cell cycle regulation were down-regulated. (D,E) *In vitro* assays of pictilisib-treated cells show significantly increased apoptosis and G0/G1 cell cycle arrest. (F,G,H) Quantification of proteins related to cell cycle regulation, apoptosis regulation, and DNA damage repair show significant differences (q-value<0.05). (I,J) *In vitro* scratch and trans-well migration assays show significantly reduced migration activity of pictilisib-treated cells. (K) Quantification of proteins related to biological adhesion shows significant differences (q-value<0.05).

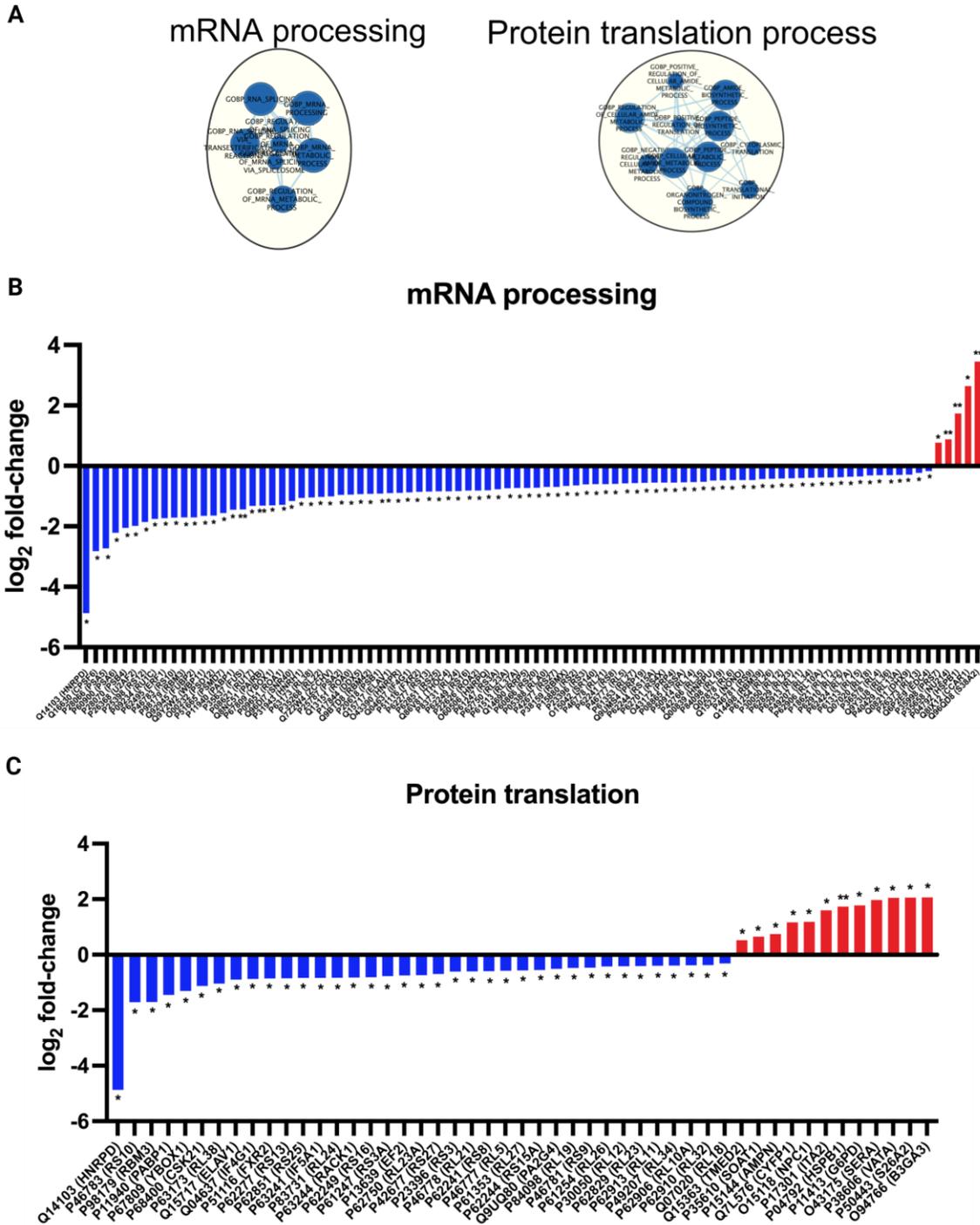


Figure 4.4 Pictilisib treatment also significantly affected pathways involving mRNA processing and protein translation in A549. (A) Based on the GSEA analysis, clusters of pathways involved in mRNA processing and protein translation are down-regulated. Most genes involved in mRNA

processing (B) and protein translation (C) are significantly under-expressed upon pictilisib treatment.

Glycoproteins with Reduced Fucosylation and Sialylation were Involved in Apoptosis, DNA Damage Repair, and Cell-Adhesion

Our glycomics results show that pictilisib treatment significantly reduced global fucosylation and sialylation of glycofocalyx N-glycans. Likewise, our proteomics results show that pictilisib treatment significantly affected adhesion, apoptotic, and cell cycle pathways. To identify which glycoproteins have reduced fucosylation and sialylation, and their involvement in these pathways, we performed quantitative site-specific glycoproteomics coupled with gene ontology analysis of pictilisib-treated cells. Glycoforms were identified after score-filtering, replicate-filtering (present in at least 2 replicates), and normalized glycopeptides per protein glycosite (**Figure 4.5A**, **Supplementary Figures 4.11** and **4.12**, and **Supplementary Table 4.5**). Normalized glycoforms were categorized based on N-glycan type - high-mannose, undecorated, fucosylated, sialylated, and sialofucosylated - and then summed for each glycosite. For example, the changes in glycoform occupancy in ANPEP, ADA10, ITGB1, and ITGA3 upon pictilisib treatment have reduced fucosylation and sialylation or sialofucosylation in specific glycosites (**Figure 4.5B**). Glycosites with reduced fucosylation, sialylation, and sialofucosylation were represented as heat maps annotated by gene ontologies of corresponding glycoproteins. Indeed, glycoproteins associated with biological adhesion and locomotion, and apoptosis, had reduced fucosylation, sialylation, and sialofucosylation (**Figure 4.5C**).

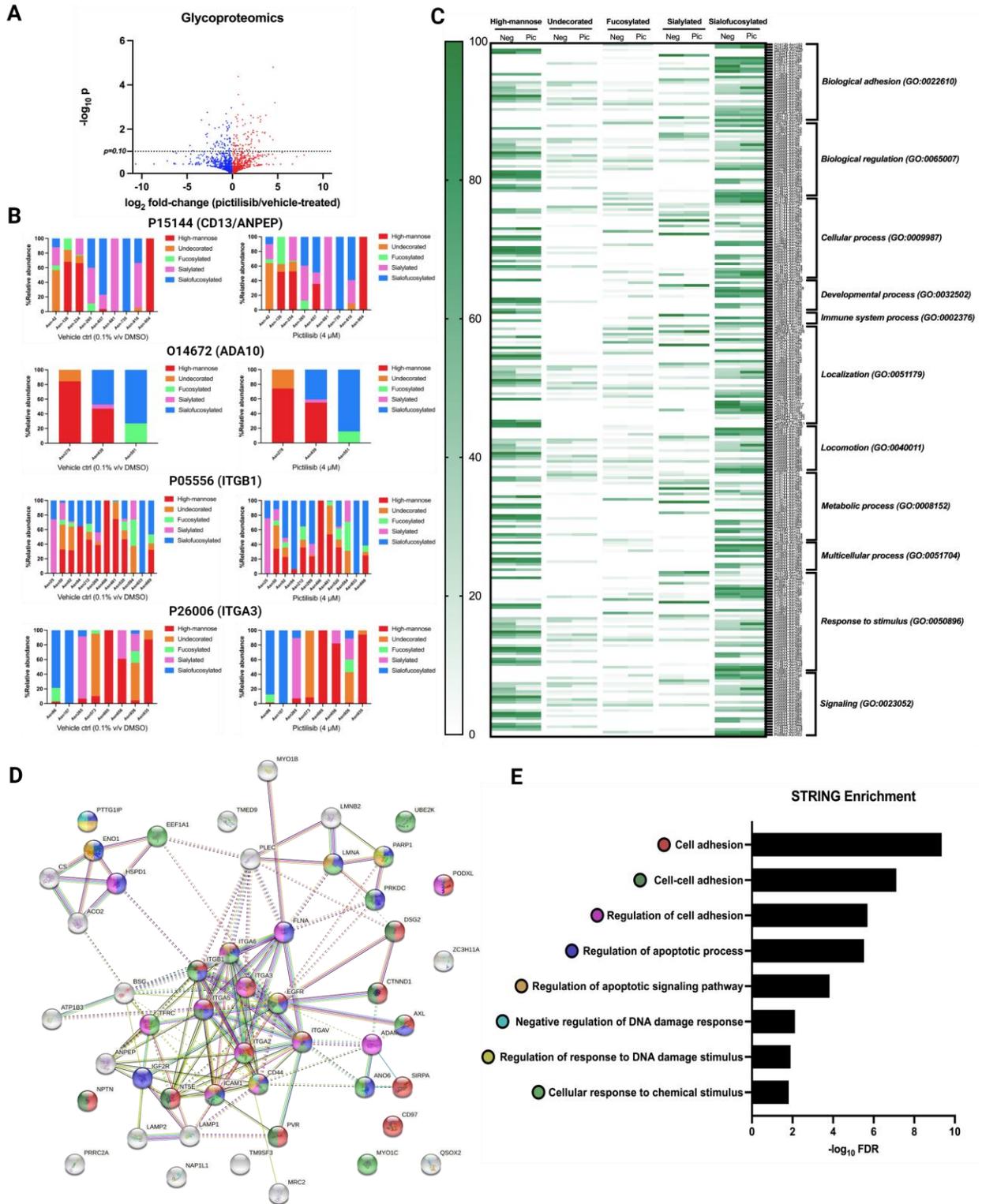


Figure 4.5 Pictilisib treatment reduced sialylation and fucosylation in specific glycoproteins. (A) Volcano plot of differentially abundant glycopeptides. (B) Site-specific glycosylation - high-

mannose, undecorated, fucosylated, sialylated, and sialofucosylated - of several glycoproteins shown to have reduced fucosylation, sialylation, or sialofucosylation upon pictilisib treatment. (C) Gene ontology analysis of proteins with reduced fucosylation, sialylation, and fucosylation show glycoproteins involved in several biological processes. (D) STRING interaction analysis shows the interaction of the glycoproteins with reduced fucosylation, sialylation, or sialofucosylation. (E) Subsequent STRING enrichment analysis shows a significant enrichment of biological processes involved in adhesion, apoptosis, response to chemicals, and DNA damage.

Interestingly, the pictilisib treatment also reduced fucosylation, sialylation, and sialofucosylation of glycoproteins involved in stimulus-response. On the other hand, glycoproteins involved in immune system response only reduced sialylation and sialofucosylation. Looking specifically at pathway effects, integrins (involved in the integrin pathway) had reduced fucosylation, sialylation, and sialofucosylation of several of their glycosites. Epidermal growth factor receptor (EGFR) pathway and ubiquitin-proteasome pathway glycoproteins had reduced sialofucosylation upon pictilisib treatment (**Supplementary Figures 4.13 and 4.14**). These glycoproteins are also shown to perform functions in binding, catalysis, regulation, signal transduction, transport, and structural support. When mapped to show protein-protein interaction network using STRING⁴⁵ (**Figure 4.5D**), enrichment analysis further confirms these glycoproteins to be significantly enriched in pathways related to cell-adhesion, apoptotic process and signaling pathways, DNA damage responses, and cellular responses to chemical stimuli (**Figure 4.5E**).

Site-specific protein glycoproteomics also allows us to investigate deeper into the molecular interactions between glycoproteins. Integrin α -5 (ITGA5) and integrin β -1 (ITGB1) are integrins involved in several biological processes, including cell-adhesion and survival. Upon pictilisib treatment, we found several glycosites in both glycoproteins that had either reduced

fucosylation, sialylation, or sialofucosylation (**Figure 4.6A and 4.6C and Supplementary Table 4.6**).

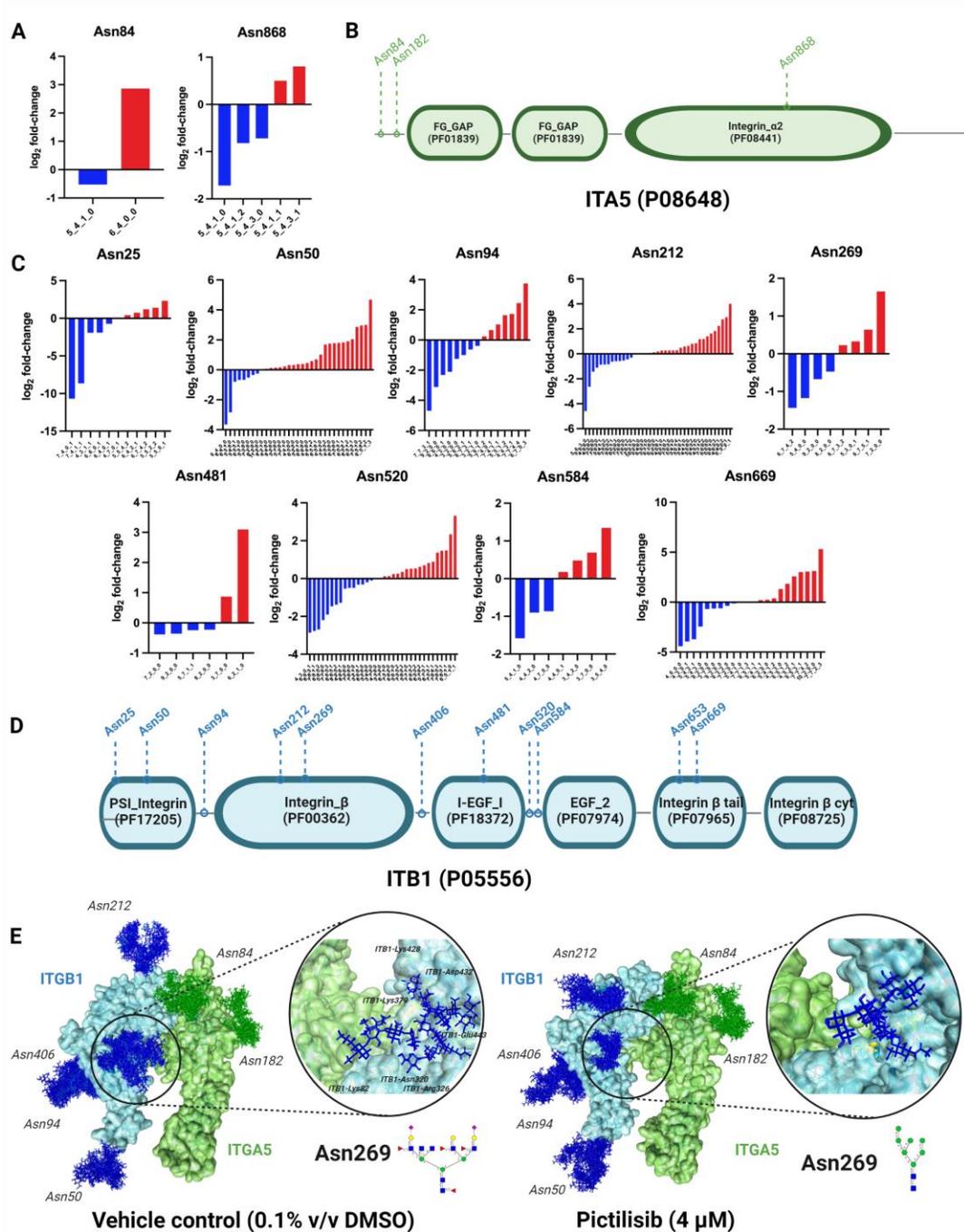


Figure 4.6 Site-specific glycosylation analysis of ITGA5-ITGB1 illustrates how specific

glycosites potentially contribute to protein interactions. (A) Site-specific glycosylation of ITGA5 glycoprotein. Bar graphs represent log₂ fold-changes in glycoform abundance upon pictilisib treatment. Glycosite Asn182 did not change glycosylation. (B) Site-specific glycosylation overlaid with protein domain information of ITGA5, annotated using PFAM (<http://pfam.xfam.org/>). (C) Site-specific glycosylation of ITGB1 glycoprotein. Bar graphs represent log₂ fold-changes in glycoform abundance upon pictilisib treatment. Glycosites Asn406 and Asn653 did not change glycosylation. The X-axis represents glycoforms, annotated as Hex_aHexNAc_bFuc_cSia_d. (D) Site-specific glycosylation overlaid with protein domain information of ITGB1, annotated using PFAM (<http://pfam.xfam.org/>). (E) 3D trajectories of ITGA5-ITGB1 glycoprotein complexes upon treatment with pictilisib. Specific glycoform structures can be seen in **Supplementary Tables 4.8** and **4.9**. Dynamics simulation of negative control (**Supplementary video 4.1**) and pictilisib-treated (**Supplementary video 4.2**) are also available.

This site-specific glycosylation information was overlaid with protein domain annotations from PFAM (<http://pfam.xfam.org/>)⁴⁶, showing how certain glycosites could potentially contribute to protein interactions (**Figure 6B** and **6D**). Further analysis through molecular dynamics shows specific glycosites to be significantly affected by sialylation. HexNAc(7)Hex(6)Fuc(4)NeuAc(2) was found to be down-regulated and HexNAc(2)Hex(7) up-regulated in ITGB1 glycosite Asn269, upon pictilisib treatment. These glycans were modeled into ITGA5 and ITGB1 complex (PDB ID: 3vi4) using CHARMM-GUI⁴⁷, then simulated over 45 ns using NAMD⁴⁸ (**Supplementary Videos 4.1** and **4.2**). Hydrogen bonding interactions were then monitored in VMD to show more residue contacts by HexNAc(7)Hex(6)Fuc(4)NeuAc(2) in the negative control compared to HexNAc(2)Hex(7) in the pictilisib-treated cells (**Figure 4.6E**).

Discussion

Aberrant glycosylation has been well-documented in cancer, with fundamental changes in the glycosylation patterns of cell-surface and excreted proteins during cancer progression. Growing evidence supported the role of glycosylation in multiple steps during tumor progression, as well as cancer cell proliferation, invasion, metastasis, and angiogenesis⁴⁹. Munkley and Elliott (2016) summarized the roles of glycosylation in cancer progression, particularly the ten hallmarks of cancer⁵⁰. To resist cell death, cancer cells must either down-regulate pro-apoptotic pathways or up-regulate anti-apoptotic/pro-survival signaling molecules. Glycans have also been shown to have a role in regulating processes that lead to cell death, such as controlling intra- and extracellular pathways that promote the initiation and execution of apoptosis⁵¹. By modifying the glycans presented on cell death receptors, cancer cells may resist apoptosis. Glycosylation can also modulate the function of death receptors of the extrinsic apoptotic pathway, Fas (CD95) and TNFR1 (tumor necrosis factor receptor 1)⁵². These glycosylations may positively regulate the apoptotic machinery. Evidence has shown the effect of glycan structures in activating invasion and metastasis pathways. Increased sialylation of surface glycoproteins has been associated with malignancy and poor patient prognosis⁵³. Increasing sialylation leads to a buildup of negative charges, physically disrupting cell-cell-adhesion and promote detachment through electrostatic repulsion⁵⁴. Over-expression of the enzyme ST6GAL1 and its glycan product sTn leads to increased migration and invasion in carcinoma⁵⁵. Glycosylation could also affect cadherin-mediated cell-adhesion. Over-expression of MGAT5, which introduces β 1-6 GlcNAc branching of N-glycans, in gastric cancer induces E-cadherin mislocalization from the cell membrane into the cytoplasm⁴⁹. This, in turn, leads to non-functioning adherens junctions, impaired cell-cell-adhesion, and signaling, and ultimately leads to enhanced metastasis⁵⁶. Once tumor cells have metastasized, invasion and colonization in distant tissue environments require upregulating

adhesion of tumor cells to endothelial cells. Glycosylation could also have a role in this – the SLe^x antigen can promote cancer cell-adhesion through interactions with selectins ⁵⁷.

The β 1,6 branching of N-glycans consists of the addition of an antenna whose first GlcNAc is β 1,6-linked to the core mannose residue ⁴⁴. This antenna is preferentially elongated by polylactosaminic sequences and frequently terminated by Lewis antigens. This increased branching expression is due to increased GlcNAcT-V activity, a glycosyltransferase enzyme encoded by the MGAT5 gene. These glycans are linked to cell surface molecules, including growth receptors PDGFR and EGFR. Galectin-3 binding to β 1,6 branched glycans regulate tumor cell motility by stimulating focal adhesion modeling, FAK and PI3K activation, local F-actin instability, and α 5 β 1 integrin translocation to fibrillar adhesions ⁵⁸. Lewis a and Lewis b antigens originated from the mono- or difucosyl substitution of type 1 chains, while Lewis x and Lewis y derive from the mono- or di-fucosyl substitution of type 2 chains. The mono-fucosyl substitution of the α 2,3-sialylated type 1 or type 2 chains lead to the formation of sialyl Lewis^a (sLe^a) and sialyl Lewis^x (sLe^x), respectively. The aberrant expression of Lewis-type antigens appears to be cancer-associated, as seen in several carcinomas, including lung cancer ⁴⁹. sLe^x and sLe^a act as ligands for E- and P-selectin cell-adhesion molecules expressed in activated endothelial cells ⁵⁹. These molecules also regulate the metastatic cascade by forming emboli of cancer cells and platelets ⁶⁰. In some cancer cell lines, the important glycoproteins carrying the sialyl Lewis antigens include the hyaluronate receptor (CD44) ⁶¹, mucin 1 (MUC1) ⁶², and lysosomal membrane glycoproteins 1 and 2 (LAMP-1 and -2). Core fucosylation is also observed in several cancers ⁴⁹. This involves the addition of α 1,6-fucose to the innermost GlcNAc residue of N-glycans through Fuc-TVIII (FUT8). Overexpression is observed in several cancers, including lung cancer ⁶³. In breast cancer, increased core fucosylation of EGFR was associated with increased dimerization and

phosphorylation, resulting in increased EGFR-mediated signaling promoting tumor growth ⁶⁴. α 2,6-sialylated lactosamine (Sia6LacNAc) is the product of β -galactoside α 2,6-sialyltransferase (ST6Gal1). The expression of this enzyme is altered in several cancers, including colon, stomach, and ovarian ⁶⁵. The Ras pathway regulates the transcription and expression of ST6Gal1, and transfectants containing ST6Gal1-expressing cells indicate increased adhesion to extracellular matrix molecules in colon ⁶⁶ and breast cancer ⁶⁷.

Conclusion

Here, we report for the first time the cytotoxic effects of an identified glycosylation inhibitor, pictilisib, on non-small cell lung carcinoma cell line (A549). By integrating network pharmacology approach and *in silico* docking methods, we were able to identify a glycosylation inhibitor that we were able to verify using mass-spectrometric glycomics approach. The compound was validated to inhibit the formation of fucosylated and sialylated N-glycans, which were primarily attached to glycoproteins involved in apoptosis, cell-adhesion, DNA damage repair, and chemical response processes. Furthermore, the compound was able to significantly affect cellular processes involved in cell cycle, apoptosis, cell-adhesion, transcription, and translation, which we were able to validate using *in vitro* biochemical assays. Finally, we modelled the differences in interactions of a model glycoprotein complex, ITGAV-ITGB1, upon changing the glycosylation by pictilisib treatment. We present for the first time how a drug, pictilisib, affects protein N-glycosylation and the pathways involving these glycoproteins through an integrated mass-spectrometric (glycomics, proteomics, and glycoproteomics) and bioinformatics pipeline

References

1. Cheng, T.-Y. D.; Cramb, S. M.; Baade, P. D.; Youlten, D. R.; Nwogu, C.; Reid, M. E. The International Epidemiology of Lung Cancer: Latest Trends, Disparities, and Tumor Characteristics. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* **2016**, *11* 10, 1653–1671. <https://doi.org/10.1016/j.jtho.2016.05.021>.
2. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **2018**, *68* 6, 394–424. <https://doi.org/10.3322/caac.21492>.
3. Meany, D. L.; Chan, D. W. Aberrant Glycosylation Associated with Enzymes as Cancer Biomarkers. *Clin. Proteomics* **2011**, *8* 1, 7. <https://doi.org/10.1186/1559-0275-8-7>.
4. Landi, M. T.; Dracheva, T.; Rotunno, M.; Figueroa, J. D.; Liu, H.; Dasgupta, A.; Mann, F. E.; Fukuoka, J.; Hames, M.; Bergen, A. W.; Murphy, S. E.; Yang, P.; Pesatori, A. C.; Consonni, D.; Bertazzi, P. A.; Wacholder, S.; Shih, J. H.; Caporaso, N. E.; Jen, J. Gene Expression Signature of Cigarette Smoking and Its Role in Lung Adenocarcinoma Development and Survival. *PLoS One* **2008**, *3* 2, e1651.
5. Ruhaak, L. R.; Taylor, S. L.; Stroble, C.; Nguyen, U. T.; Parker, E. A.; Song, T.; Lebrilla, C. B.; Rom, W. N.; Pass, H.; Kim, K.; Kelly, K.; Miyamoto, S. Differential N-Glycosylation Patterns in Lung Adenocarcinoma Tissue. *J. Proteome Res.* **2015**, *14* 11, 4538–4549. <https://doi.org/10.1021/acs.jproteome.5b00255>.
6. Ogawa, J.; Inoue, H.; Koide, S. Expression of Alpha-1,3-Fucosyltransferase Type IV and VII Genes Is Related to Poor Prognosis in Lung Cancer. *Cancer Res.* **1996**, *56* 2, 325–329.
7. Vajaria, B. N.; Patel, P. S. Glycosylation: A Hallmark of Cancer? *Glycoconj. J.* **2017**, *34* 2, 147–156. <https://doi.org/10.1007/s10719-016-9755-2>.
8. Handerson, T.; Camp, R.; Harigopal, M.; Rimm, D.; Pawelek, J. Beta1,6-Branched Oligosaccharides Are Increased in Lymph Node Metastases and Predict Poor Outcome in Breast Carcinoma. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **2005**, *11* 8, 2969–2973. <https://doi.org/10.1158/1078-0432.CCR-04-2211>.
9. Recchi, M. A.; Hebbbar, M.; Hornez, L.; Harduin-Lepers, A.; Peyrat, J. P.; Delannoy, P. Multiplex Reverse Transcription Polymerase Chain Reaction Assessment of Sialyltransferase Expression in Human Breast Cancer. *Cancer Res.* **1998**, *58* 18, 4066–4070.
10. Burchell, J.; Poulsom, R.; Hanby, A.; Whitehouse, C.; Cooper, L.; Clausen, H.; Miles, D.; Taylor-Papadimitriou, J. An Alpha2,3 Sialyltransferase ST3Gal I. Is Elevated in Primary Breast Carcinomas. *Glycobiology* **1999**, *9* 12, 1307–1311. <https://doi.org/10.1093/glycob/9.12.1307>.
11. Picco, G.; Julien, S.; Brockhausen, I.; Beatson, R.; Antonopoulos, A.; Haslam, S.; Mandel, U.; Dell, A.; Pinder, S.; Taylor-Papadimitriou, J.; Burchell, J. Over-Expression of ST3Gal-I Promotes Mammary Tumorigenesis. *Glycobiology* **2010**, *20* 10, 1241–1250. <https://doi.org/10.1093/glycob/cwq085>.
12. Schneider, F.; Kemmner, W.; Haensch, W.; Franke, G.; Gretschel, S.; Karsten, U.; Schlag, P. M. Overexpression of Sialyltransferase CMP-Sialic Acid:Galbeta1,3GalNAc-R Alpha6-Sialyltransferase Is Related to Poor Patient Survival in Human Colorectal Carcinomas. *Cancer Res.* **2001**, *61* 11, 4605–4611.
13. Esko, J. D.; Bertozzi, C. R. Chemical Tools for Inhibiting Glycosylation.; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; *Essentials of Glycobiology*. 2nd ed. Cold Spring Harbor, Cold Spring Harbor NY., 2009.

14. Hiss, D. C.; Gabriels, G. A.; Folb, P. I. Combination of Tunicamycin with Anticancer Drugs Synergistically Enhances Their Toxicity in Multidrug-Resistant Human Ovarian Cystadenocarcinoma Cells. *Cancer Cell Int.* **2007**, *7*, 5. <https://doi.org/10.1186/1475-2867-7-5>.
15. Dong, Y.; Yin, S.; Jiang, C.; Luo, X.; Guo, X.; Zhao, C.; Fan, L.; Meng, Y.; Lu, J.; Song, X.; Zhang, X.; Chen, N.; Hu, H. Involvement of Autophagy Induction in Penta-1,2,3,4,6-O-Galloyl- β -D-Glucose-Induced Senescence-like Growth Arrest in Human Cancer Cells. *Autophagy* **2014**, *10* 2, 296–310. <https://doi.org/10.4161/auto.27210>.
16. Davis, A. P.; Grondin, C. J.; Johnson, R. J.; Sciaky, D.; Wieggers, J.; Wieggers, T. C.; Mattingly, C. J. Comparative Toxicogenomics Database CTD.: Update 2021. *Nucleic Acids Res.* **2021**, *49* D1, D1138–D1143. <https://doi.org/10.1093/nar/gkaa891>.
17. Szklarczyk, D.; Santos, A.; von Mering, C.; Jensen, L. J.; Bork, P.; Kuhn, M. STITCH 5: Augmenting Protein-Chemical Interaction Networks with Tissue and Affinity Data. *Nucleic Acids Res.* **2016**, *44* D1, D380–D384. <https://doi.org/10.1093/nar/gkv1277>.
18. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T. I.; Nudel, R.; Lieder, I.; Mazor, Y.; Kaplan, S.; Dahary, D.; Warshawsky, D.; Guan-Golan, Y.; Kohn, A.; Rappaport, N.; Safran, M.; Lancet, D. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **2016**, *54*, 1.30.1-1.30.33. <https://doi.org/10.1002/cpbi.5>.
19. Freshour, S. L.; Kiwala, S.; Cotto, K. C.; Coffman, A. C.; McMichael, J. F.; Song, J. J.; Griffith, M.; Griffith, O. L.; Wagner, A. H. Integration of the Drug-Gene Interaction Database DGIdb 4.0. with Open Crowdsourced Efforts. *Nucleic Acids Res.* **2021**, *49* D1, D1144–D1151. <https://doi.org/10.1093/nar/gkaa1084>.
20. Burley, S. K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G. V.; Christie, C. H.; Dalenberg, K.; Di Costanzo, L.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ganesan, S.; Goodsell, D. S.; Ghosh, S.; Green, R. K.; Guranović, V.; Guzenko, D.; Hudson, B. P.; Lawson, C. L.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Persikova, I.; Randle, C.; Rose, A.; Rose, Y.; Sali, A.; Segura, J.; Sekharan, M.; Shao, C.; Tao, Y.-P.; Voigt, M.; Westbrook, J. D.; Young, J. Y.; Zardecki, C.; Zhuravleva, M. RCSB Protein Data Bank: Powerful New Tools for Exploring 3D Structures of Biological Macromolecules for Basic and Applied Research and Education in Fundamental Biology, Biomedicine, Biotechnology, Bioengineering and Energy Sciences. *Nucleic Acids Res.* **2021**, *49* D1, D437–D451. <https://doi.org/10.1093/nar/gkaa1038>.
21. Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* **2018**, *46* D1, D1074–D1082. <https://doi.org/10.1093/nar/gkx1037>.
22. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* 11, 2498–2504. <https://doi.org/10.1101/gr.1239303>.
23. Dallakyan, S.; Olson, A. J. Small-Molecule Library Screening by Docking with PyRx BT - Chemical Biology: Methods and Protocols; Hempel, J. E., Williams, C. H., Hong, C. C., Eds.; Springer New York: New York, NY, 2015; pp 243–250. https://doi.org/10.1007/978-1-4939-2269-7_19.
24. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J.*

- Am. Chem. Soc.* **1992**, *114* 25, 10024–10035. <https://doi.org/10.1021/ja00051a040>.
25. O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3* 1, 33. <https://doi.org/10.1186/1758-2946-3-33>.
26. Nagae, M.; Kizuka, Y.; Mihara, E.; Kitago, Y.; Hanashima, S.; Ito, Y.; Takagi, J.; Taniguchi, N.; Yamaguchi, Y. Structure and Mechanism of Cancer-Associated N-Acetylglucosaminyltransferase-V. *Nat. Commun.* **2018**, *9* 1, 3380. <https://doi.org/10.1038/s41467-018-05931-w>.
27. Kuhn, B.; Benz, J.; Greif, M.; Engel, A. M.; Sobek, H.; Rudolph, M. G. The Structure of Human α -2,6-Sialyltransferase Reveals the Binding Mode of Complex Glycans. *Acta Crystallogr. D. Biol. Crystallogr.* **2013**, *69* Pt 9, 1826–1838. <https://doi.org/10.1107/S0907444913015412>.
28. Ihara, H.; Ikeda, Y.; Toma, S.; Wang, X.; Suzuki, T.; Gu, J.; Miyoshi, E.; Tsukihara, T.; Honke, K.; Matsumoto, A.; Nakagawa, A.; Taniguchi, N. Crystal Structure of Mammalian A1,6-Fucosyltransferase, FUT8. *Glycobiology* **2007**, *17* 5, 455–466. <https://doi.org/10.1093/glycob/cwl079>.
29. Lira-Navarrete, E.; Valero-González, J.; Villanueva, R.; Martínez-Júlvez, M.; Tejero, T.; Merino, P.; Panjikar, S.; Hurtado-Guerrero, R. Structural Insights into the Mechanism of Protein O-Fucosylation. *PLoS One* **2011**, *6* 9, e25365.
30. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46* W1, W296–W303. <https://doi.org/10.1093/nar/gky427>.
31. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* 13, 1605–1612. <https://doi.org/10.1002/jcc.20084>.
32. Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* 2, 455–461. <https://doi.org/10.1002/jcc.21334>.
33. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9* 7, 671–675. <https://doi.org/10.1038/nmeth.2089>.
34. 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>.
35. Stanley, P.; Taniguchi, N.; Aebi, M. N-Glycans.; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Essentials of Glycobiology. 2nd ed. Cold Spring Harbor, Cold Spring Harbor NY., 2015; pp 99–111. <https://doi.org/10.1101/glycobiology.3e.009>.
36. The Uniprot Consortium. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49* D1, D480–D489. <https://doi.org/10.1093/nar/gkaa1100>.
37. Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. A.; Paulovich, A.; Pomeroy, S. L.; Golub, T. R.; Lander, E. S.; Mesirov, J. P. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* 43, 15545–15550. <https://doi.org/10.1073/pnas.0506580102>.

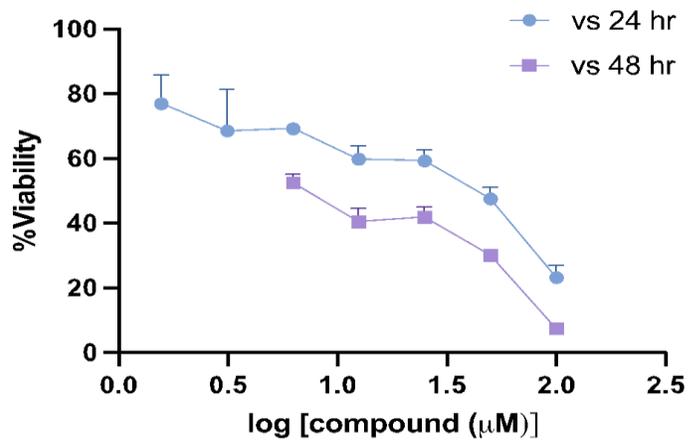
38. Raudvere, U.; Kolberg, L.; Kuzmin, I.; Arak, T.; Adler, P.; Peterson, H.; Vilo, J. G:Profiler: A Web Server for Functional Enrichment Analysis and Conversions of Gene Lists 2019 Update. *Nucleic Acids Res.* **2019**, *47* W1, W191–W198. <https://doi.org/10.1093/nar/gkz369>.
39. Park, S.-J.; Lee, J.; Qi, Y.; Kern, N. R.; Lee, H. S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im, W. CHARMM-GUI Glycan Modeler for Modeling and Simulation of Carbohydrates and Glycoconjugates. *Glycobiology* **2019**, *29* 4, 320–331. <https://doi.org/10.1093/glycob/cwz003>.
40. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* 1, 71–73. <https://doi.org/10.1038/nmeth.4067>.
41. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* 1, 33–38. <https://doi.org/https://doi.org/10.1016/0263-785596.00018-5>.
42. Manabe, Y.; Kasahara, S.; Takakura, Y.; Yang, X.; Takamatsu, S.; Kamada, Y.; Miyoshi, E.; Yoshidome, D.; Fukase, K. Development of A1,6-Fucosyltransferase Inhibitors through the Diversity-Oriented Syntheses of GDP-Fucose Mimics Using the Coupling between Alkyne and Sulfonyl Azide. *Bioorg. Med. Chem.* **2017**, *25* 11, 2844–2850. <https://doi.org/10.1016/j.bmc.2017.02.036>.
43. García-García, A.; Ceballos-Laita, L.; Serna, S.; Artschwager, R.; Reichardt, N. C.; Corzana, F.; Hurtado-Guerrero, R. Structural Basis for Substrate Specificity and Catalysis of A1,6-Fucosyltransferase. *Nat. Commun.* **2020**, *11* 1, 973. <https://doi.org/10.1038/s41467-020-14794-z>.
44. Dall’Olio, F.; Malagolini, N.; Trinchera, M.; Chiricolo, M. Mechanisms of Cancer-Associated Glycosylation Changes. *Front. Biosci. Landmark Ed.* **2012**, *17*, 670–699. <https://doi.org/10.2741/3951>.
45. Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.; Bork, P.; Jensen, L. J.; Mering, C. von. STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res.* **2019**, *47* D1, D607–D613. <https://doi.org/10.1093/nar/gky1131>.
46. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D. The Pfam Protein Families Database in 2019. *Nucleic Acids Res.* **2019**, *47* D1, D427–D432. <https://doi.org/10.1093/nar/gky995>.
47. Mallajosyula, S. S.; Jo, S.; Im, W.; MacKerell Jr, A. D. Molecular Dynamics Simulations of Glycoproteins Using CHARMM. *Methods Mol. Biol.* **2015**, *1273*, 407–429. https://doi.org/10.1007/978-1-4939-2343-4_25.
48. Acun, B.; Hardy, D. J.; Kale, L. V.; Li, K.; Phillips, J. C.; Stone, J. E. Scalable Molecular Dynamics with NAMD on the Summit System. *IBM J. Res. Dev.* **2018**, *62* 6.. <https://doi.org/10.1147/JRD.2018.2888986>.
49. Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat. Rev. Cancer* **2015**, *15* 9, 540–555. <https://doi.org/10.1038/nrc3982>.
50. Munkley, J.; Elliott, D. J. Hallmarks of Glycosylation in Cancer. *Oncotarget* **2016**, *7* 23, 35478–35489. <https://doi.org/10.18632/oncotarget.8155>.
51. Lichtenstein, R. G.; Rabinovich, G. A. Glycobiology of Cell Death: When Glycans and Lectins Govern Cell Fate. *Cell Death Differ.* **2013**, *20* 8, 976–986. <https://doi.org/10.1038/cdd.2013.50>.

52. Shatnyeva, O. M.; Kubarenko, A. V.; Weber, C. E. M.; Pappa, A.; Schwartz-Albiez, R.; Weber, A. N. R.; Krammer, P. H.; Lavrik, I. N. Modulation of the CD95-Induced Apoptosis: The Role of CD95 N-Glycosylation. *PLoS One* **2011**, *6* 5, e19927. <https://doi.org/10.1371/journal.pone.0019927>.
53. Amado, M.; Carneiro, F.; Seixas, M.; Clausen, H.; Sobrinho-Simões, M. Dimeric Sialyl-Lex. Expression in Gastric Carcinoma Correlates with Venous Invasion and Poor Outcome. *Gastroenterology* **1998**, *114* 3, 462–470. <https://doi.org/10.1016/s0016-508598.70529-3>.
54. Seidenfaden, R.; Krauter, A.; Schertzinger, F.; Gerardy-Schahn, R.; Hildebrandt, H. Polysialic Acid Directs Tumor Cell Growth by Controlling Heterophilic Neural Cell Adhesion Molecule Interactions. *Mol. Cell. Biol.* **2003**, *23* 16, 5908–5918. <https://doi.org/10.1128/MCB.23.16.5908-5918.2003>.
55. Munkley, J. The Role of Sialyl-Tn in Cancer. *Int. J. Mol. Sci.* **2016**, *17* 3, 275. <https://doi.org/10.3390/ijms17030275>.
56. Lucena, M. C.; Carvalho-Cruz, P.; Donadio, J. L.; Oliveira, I. A.; de Queiroz, R. M.; Marinho-Carvalho, M. M.; de Paula, I. F.; Gondim, K. C.; McComb, M. E.; Costello, C. E.; Whelan, S. A.; Todeschini, A. R.; Dias, W. B. Epithelial Mesenchymal Transition Induces Aberrant Glycosylation Through Hexosamine Biosynthetic Pathway Activation. *J. Biol. Chem.* **2016**, *17*;291(25):12917-29 <https://doi.org/10.1074/jbc.M116.729236>.
57. Numahata, K.; Satoh, M.; Handa, K.; Saito, S.; Ohyama, C.; Ito, A.; Takahashi, T.; Hoshi, S.; Orikasa, S.; Hakomori, S. Sialosyl-Lex. Expression Defines Invasive and Metastatic Properties of Bladder Carcinoma. *Cancer* **2002**, *94* 3, 673–685. <https://doi.org/10.1002/cncr.10268>.
58. Lagana, A.; Goetz, J. G.; Cheung, P.; Raz, A.; Dennis, J. W.; Nabi, I. R. Galectin Binding to Mgat5-Modified N-Glycans Regulates Fibronectin Matrix Remodeling in Tumor Cells. *Mol. Cell. Biol.* **2006**, *26* 8, 3181–3193. <https://doi.org/10.1128/MCB.26.8.3181-3193.2006>.
59. Mannori, G.; Crottet, P.; Cecconi, O.; Hanasaki, K.; Aruffo, A.; Nelson, R. M.; Varki, A.; Bevilacqua, M. P. Differential Colon Cancer Cell-adhesion to E-, P-, and L-Selectin: Role of Mucin-Type Glycoproteins. *Cancer Res.* **1995**, *55* 19, 4425–4431.
60. Witz, I. P. The Selectin-Selectin Ligand Axis in Tumor Progression. *Cancer Metastasis Rev.* **2008**, *27* 1, 19–30. <https://doi.org/10.1007/s10555-007-9101-z>.
61. Lim, S. D.; Young, A. N.; Paner, G. P.; Amin, M. B. Prognostic Role of CD44 Cell Adhesion Molecule Expression in Primary and Metastatic Renal Cell Carcinoma: A Clinicopathologic Study of 125 Cases. *Virchows Arch.* **2008**, *452* 1, 49–55. <https://doi.org/10.1007/s00428-007-0530-4>.
62. Tomlinson, J.; Wang, J. L.; Barsky, S. H.; Lee, M. C.; Bischoff, J.; Nguyen, M. Human Colon Cancer Cells Express Multiple Glycoprotein Ligands for E-Selectin. *Int. J. Oncol.* **2000**, *16* 2, 347–353. <https://doi.org/10.3892/ijo.16.2.347>.
63. Liu, Y.-C.; Yen, H.-Y.; Chen, C.-Y.; Chen, C.-H.; Cheng, P.-F.; Juan, Y.-H.; Chen, C.-H.; Khoo, K.-H.; Yu, C.-J.; Yang, P.-C.; Hsu, T.-L.; Wong, C.-H. Sialylation and Fucosylation of Epidermal Growth Factor Receptor Suppress Its Dimerization and Activation in Lung Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* 28, 11332–11337. <https://doi.org/10.1073/pnas.1107385108>.
64. Potapenko, I. O.; Haakensen, V. D.; Lüders, T.; Helland, A.; Bukholm, I.; Sørli, T.; Kristensen, V. N.; Lingjaerde, O. C.; Børresen-Dale, A.-L. Glycan Gene Expression Signatures in Normal and Malignant Breast Tissue; Possible Role in Diagnosis and Progression. *Mol. Oncol.* **2010**, *4* 2, 98–118. <https://doi.org/10.1016/j.molonc.2009.12.001>.

65. Dall'Olio, F.; Chiricolo, M. Sialyltransferases in Cancer. *Glycoconj. J.* **2001**, *18* 11–12, 841–850. <https://doi.org/10.1023/a:1022288022969>.
66. Seales, E. C.; Shaikh, F. M.; Woodard-Grice, A. V; Aggarwal, P.; McBrayer, A. C.; Hennessy, K. M.; Bellis, S. L. A Protein Kinase C/Ras/ERK Signaling Pathway Activates Myeloid Fibronectin Receptors by Altering Beta1 Integrin Sialylation. *J. Biol. Chem.* **2005**, *280* 45, 37610–37615. <https://doi.org/10.1074/jbc.M508476200>.
67. Lin, S.; Kemmner, W.; Grigull, S.; Schlag, P. M. Cell Surface Alpha 2,6 Sialylation Affects Adhesion of Breast Carcinoma Cells. *Exp. Cell Res.* **2002**, *276* 1, 101–110. <https://doi.org/10.1006/excr.2002.5521>.

Supplementary Figures

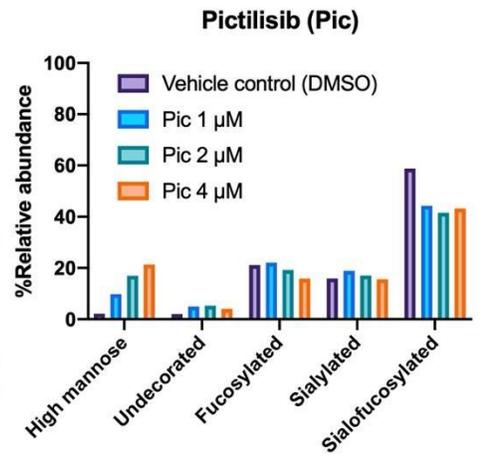
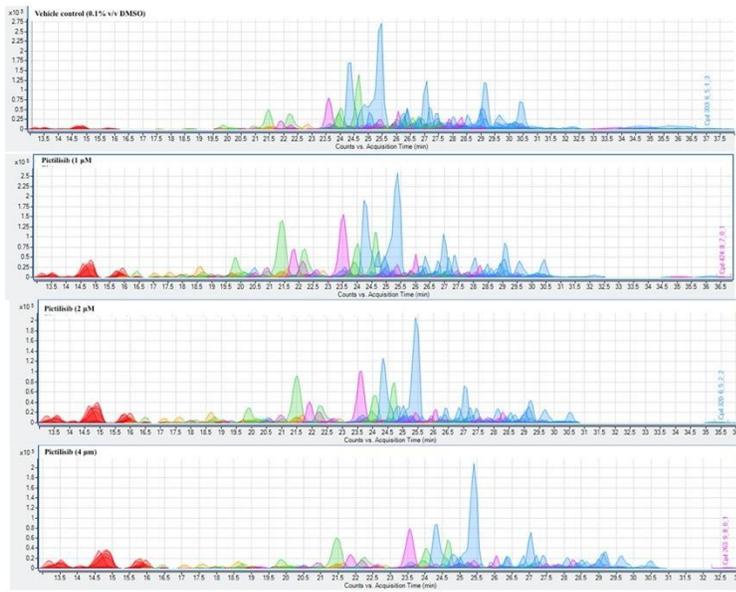
Dose-response of Pictilisib vs. A549 (24 & 48 hr)



IC₅₀-24 hr (Pic): 24.68 μM (95%CI: 16.28–37.41 μM)

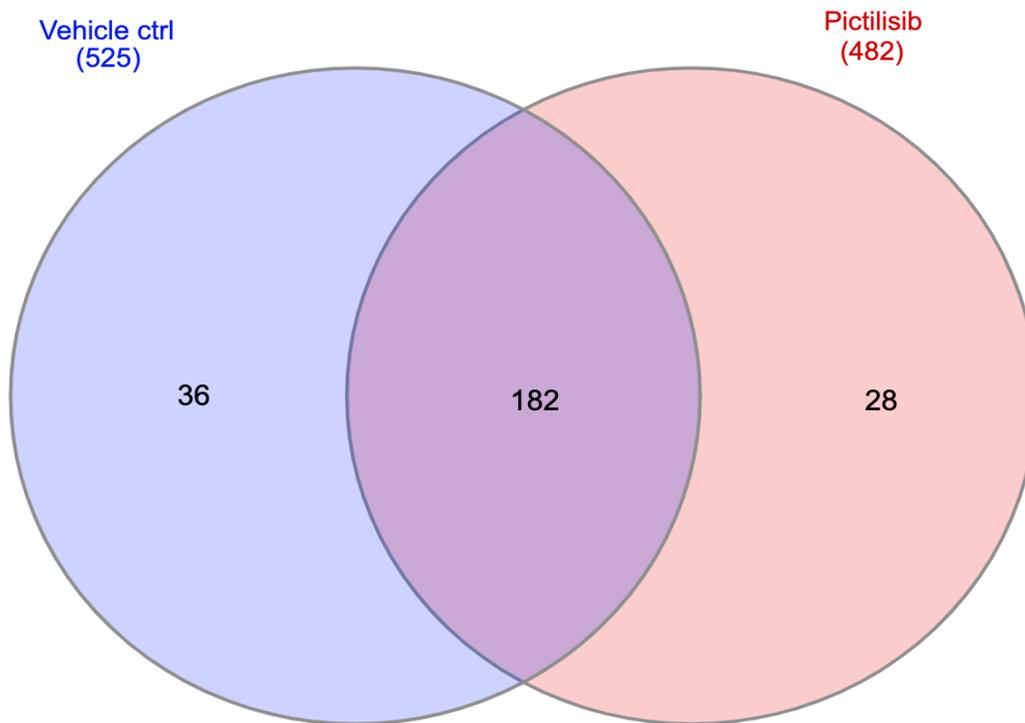
IC₅₀-48 hr (Pic): 10.84 μM (95%CI: 7.92–14.64 μM)

Supplementary figure 4.1 24- and 48-hr dose-response toxicity of pictilisib.

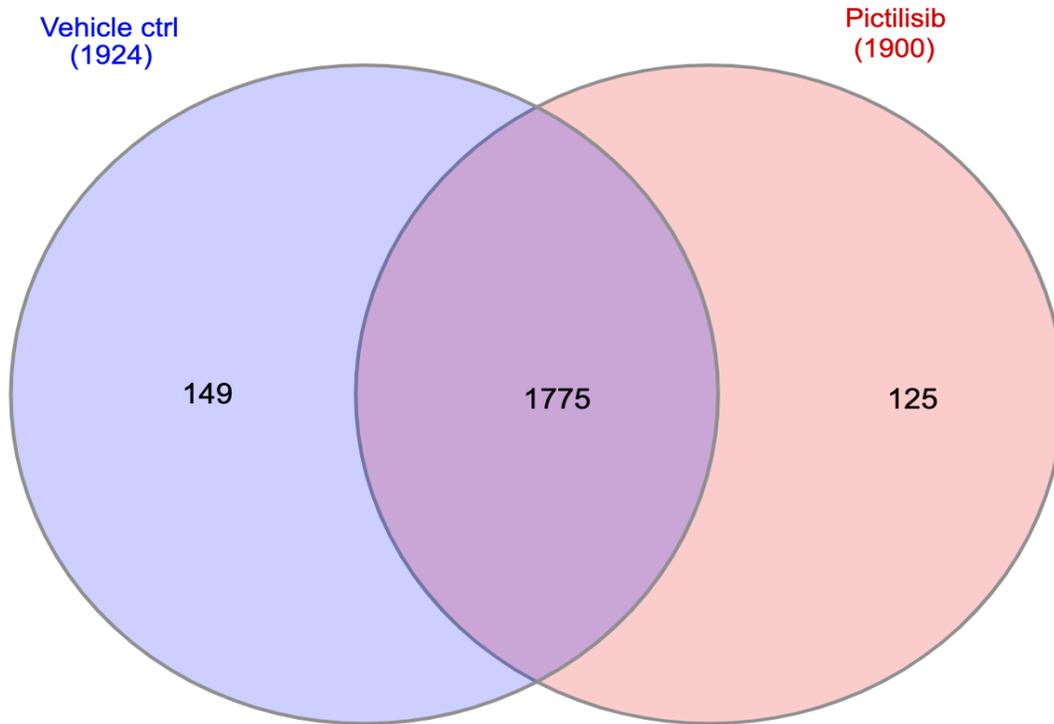


High mannose Undecorated Fucosylated Sialylated
Sialofucosylated

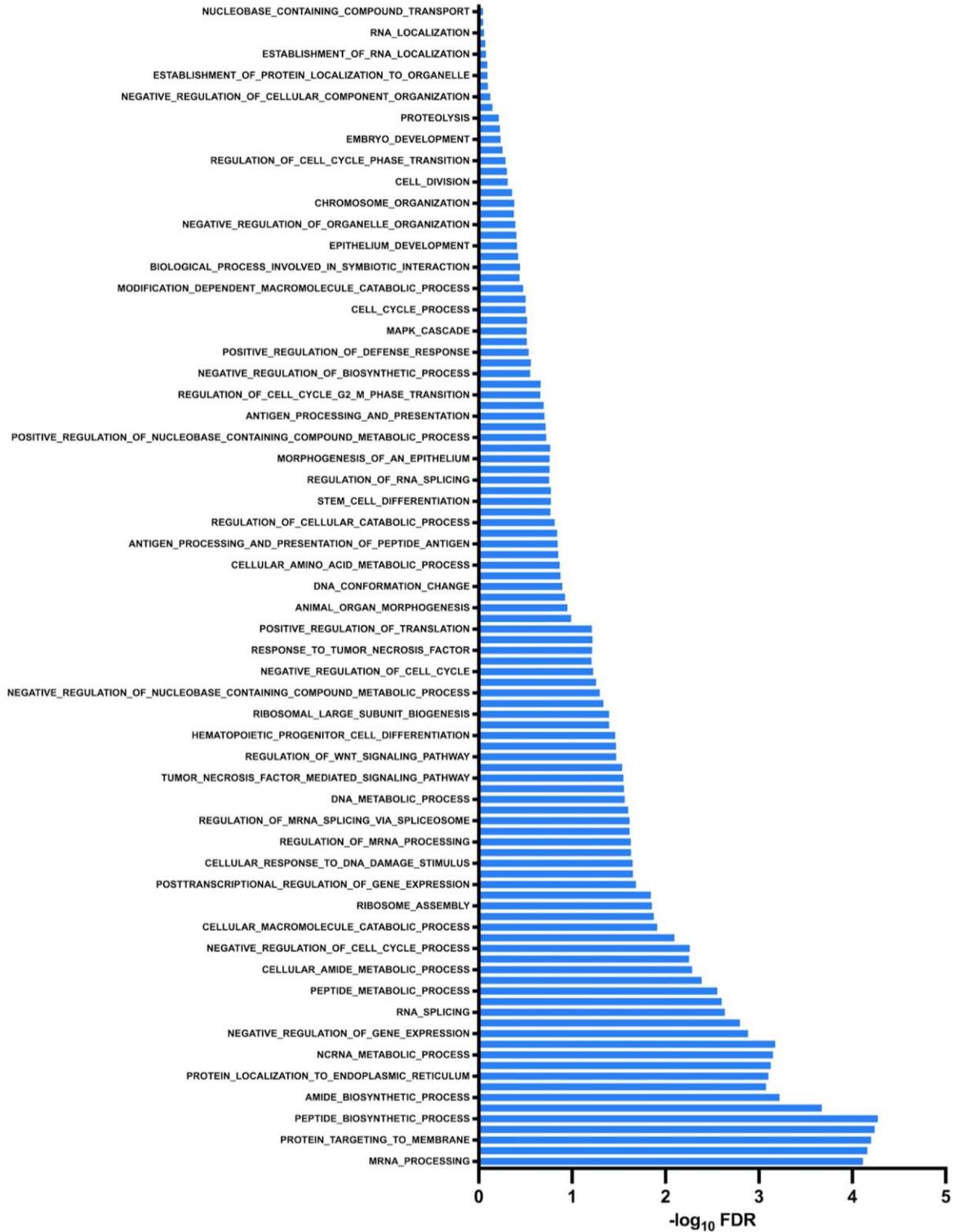
Supplementary Figure 4.2 Dose-response glycomics assay of pictilisib.



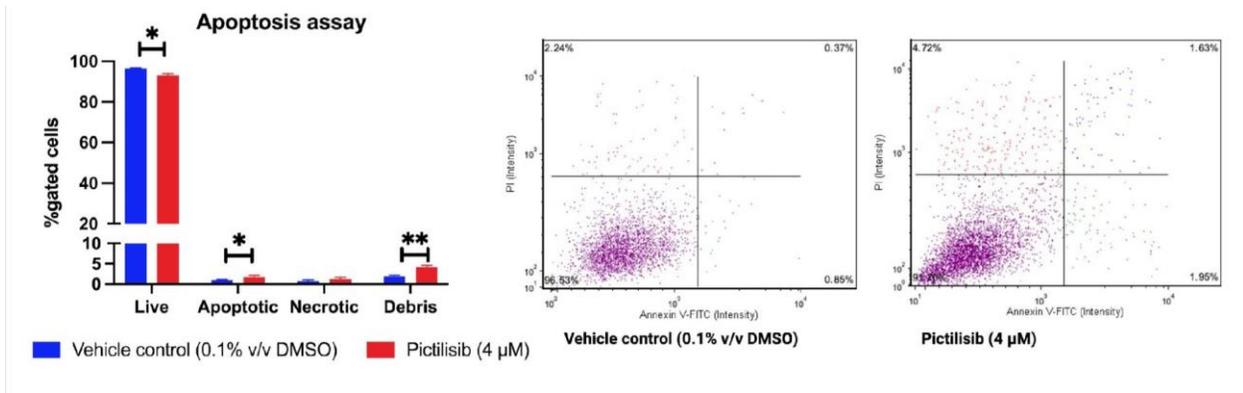
Supplementary Figure 4.3 Reproducibility of glycomics results between groups.



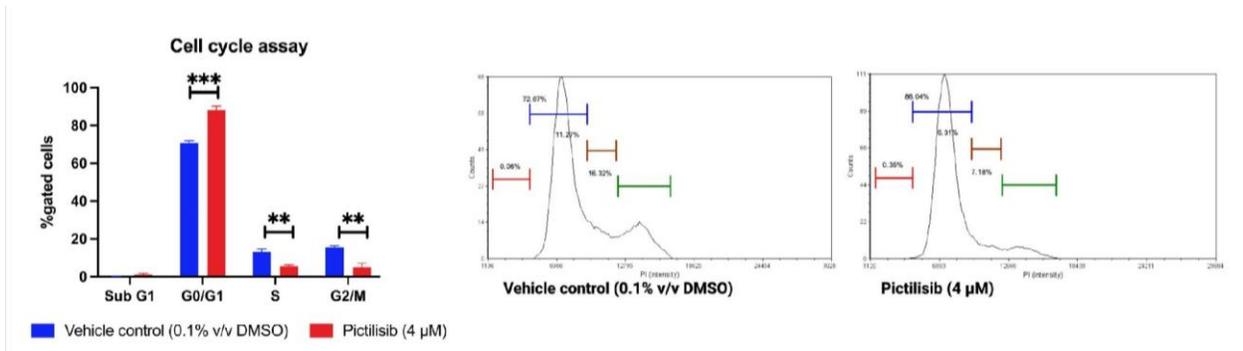
Supplementary Figure 4.4 Reproducibility of proteomics results between groups.



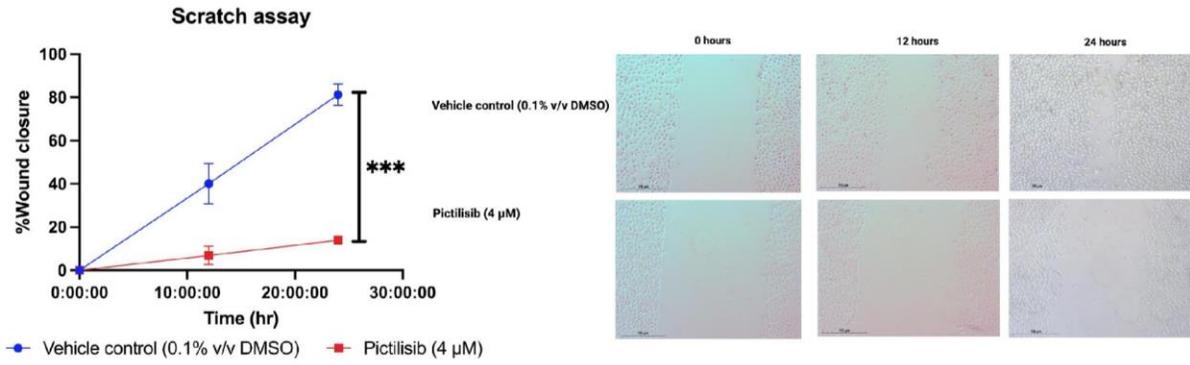
Supplementary Figure 4.5 $-\log_{10}$ FDR value results of under-expressed proteins by GSEA using Biological Process Gene Ontology set.



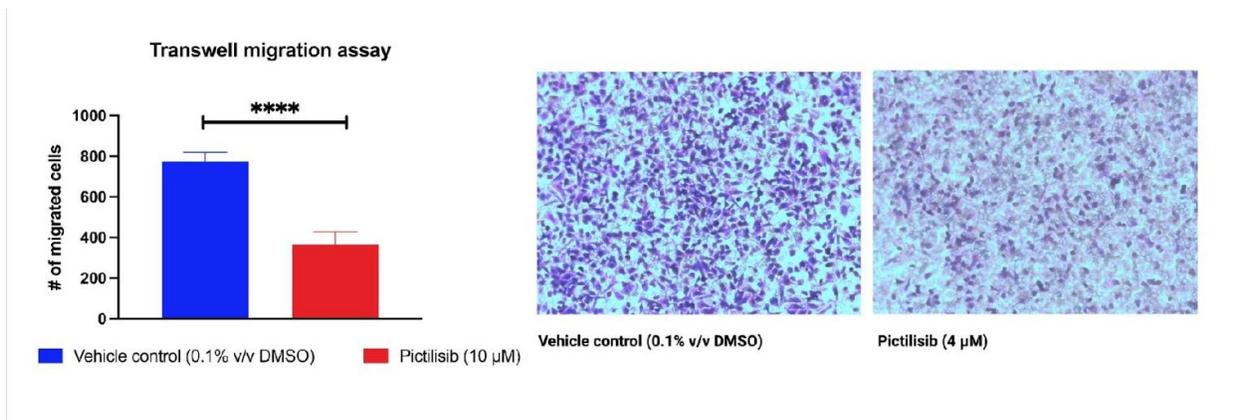
Supplementary Figure 4.7 Apoptosis assay results of pictilisib treatment.



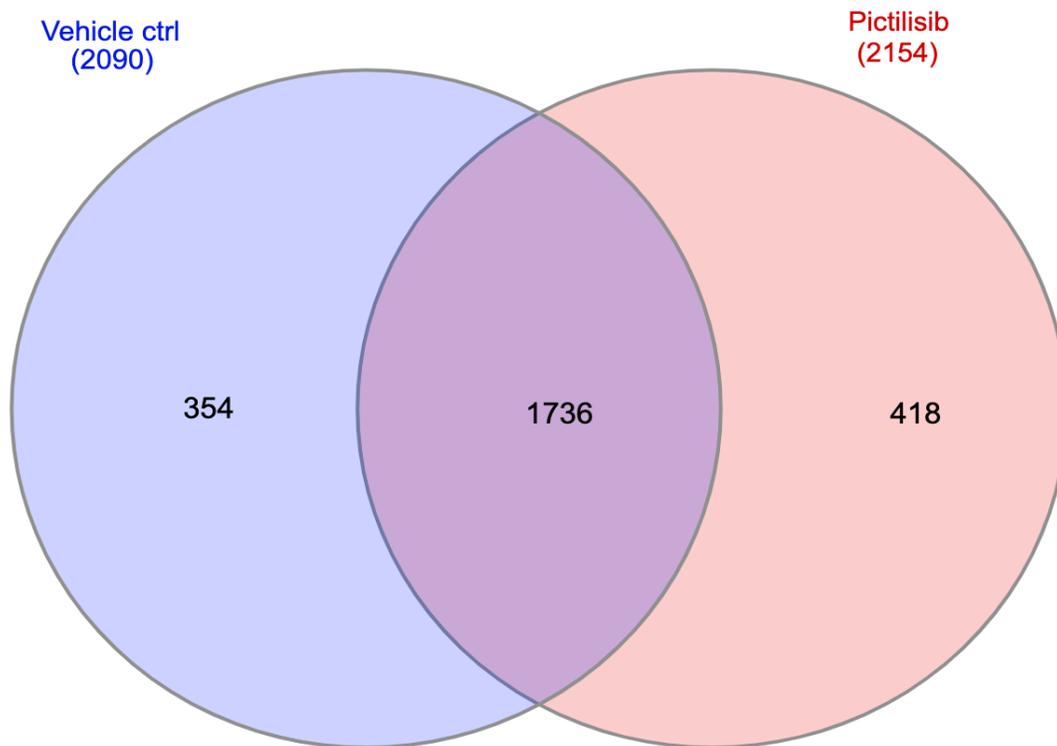
Supplementary Figure 4.8 Cell cycle assay results.



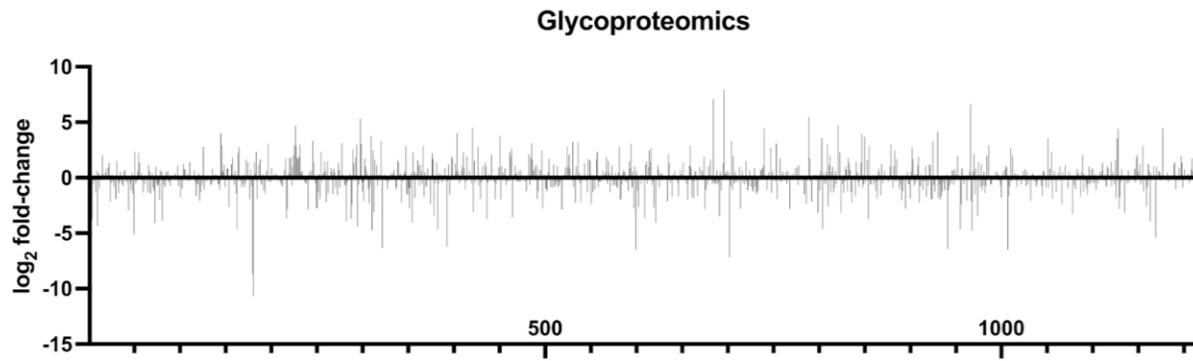
Supplementary Figure 4.9 Scratch assay results of pictilisib-treatment.



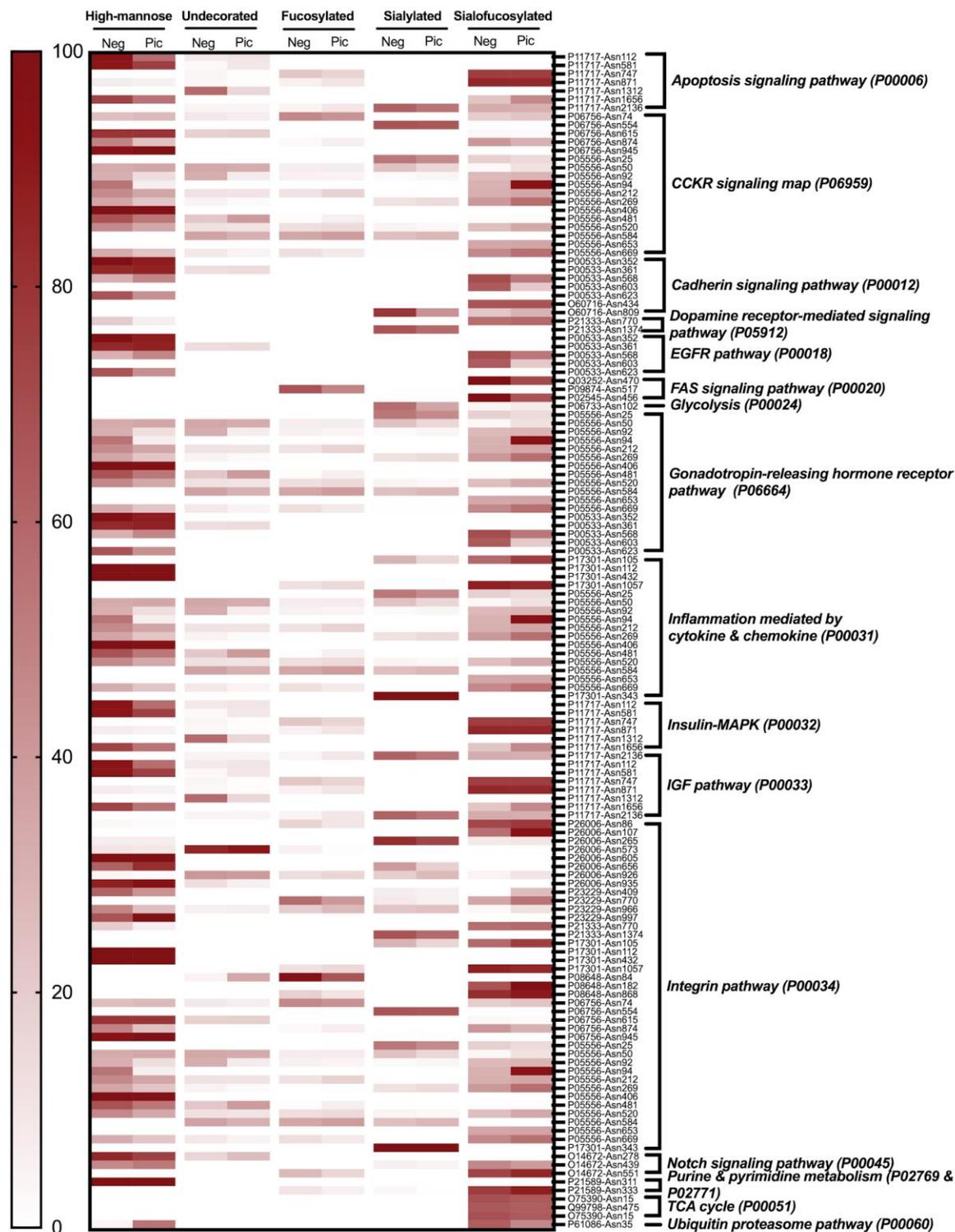
Supplementary Figure 4.10 Trans-well migration assay results of pictilisib-treatment.



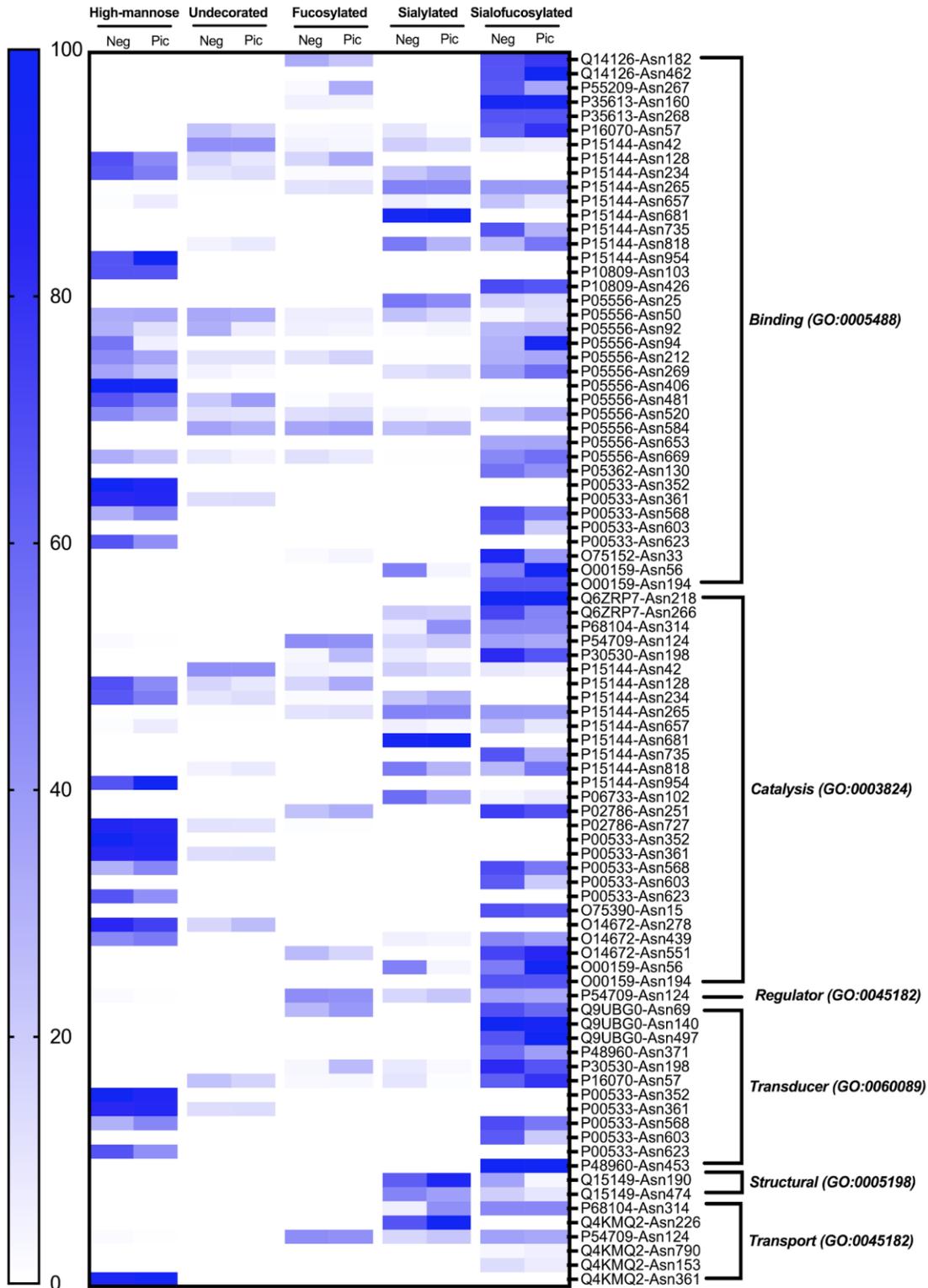
Supplementary Figure 4.11 Reproducibility of glycoproteomics results between groups.



Supplementary Figure 4.12 Variation of protein glycosylation across all glycosites.



Supplementary Figure 4.13 Pathways of proteins with reduced fucosylation, sialylation, or sialofucosylation.



Supplementary Figure 4.14 Molecular function of proteins with reduced fucosylation, sialylation, or sialofucosylation.