

Medical University of South Carolina

**MEDICA**

---

MUSC Theses and Dissertations

---

2020

# Determination of N-Linked Glycosylation Changes in Hepatocellular Carcinoma and the Associated Glycoproteins for Enhanced Biomarker Discovery and Therapeutic Targets

Connor Allen West

*Medical University of South Carolina*

Follow this and additional works at: <https://medica-musc.researchcommons.org/theses>

---

## Recommended Citation

West, Connor Allen, "Determination of N-Linked Glycosylation Changes in Hepatocellular Carcinoma and the Associated Glycoproteins for Enhanced Biomarker Discovery and Therapeutic Targets" (2020). *MUSC Theses and Dissertations*. 62.

<https://medica-musc.researchcommons.org/theses/62>

This Dissertation is brought to you for free and open access by MEDICA. It has been accepted for inclusion in MUSC Theses and Dissertations by an authorized administrator of MEDICA. For more information, please contact [medica@musc.edu](mailto:medica@musc.edu).

Determination of N-Linked Glycosylation Changes in  
Hepatocellular Carcinoma and the Associated Glycoproteins for  
Enhanced Biomarker Discovery and Therapeutic Targets

Connor Allen West

A dissertation submitted to the faculty of the Medical University of South Carolina  
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in  
the College of Graduate Studies

Department of Cell and Molecular Pharmacology and Experimental Therapeutics  
2020

Approved by:

Chairman, Advisory Committee

---

Richard Drake, PhD

---

Anand Mehta, D.Phil

---

Peggi Angel, PhD

---

Patrick Woster, PhD

---

Robin Muise-Helmericks, PhD

## TABLE OF CONTENTS

ABSTRACT	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
ACKNOWLEDGEMENTS	ix
Chapter 1: Literature Review	1
1.1. Introduction to Hepatocellular Carcinoma	2
1.1.1. Symptoms, Mortality, and Standard of Care	2
1.1.2. Liver Cancer Subtypes	4
1.1.3. Disease Etiology	7
1.1.3.1. Risk Factors	7
1.1.3.2. Common Signaling Pathways	8
1.1.3.3. Origin of Disease Initiation	9
1.1.4. Current Detection and Diagnosis of Liver Cancer	11
1.1.4.1. Clinical Detection Approaches	11
1.2. Glycosylation	13
1.2.1. General Principles of Glycosylation	13
1.2.1.1. Mechanism of Protein Glycosylation	14
1.2.1.2. Physiological and Pathological Roles of Glycosylation	20
1.2.1.2.1. N-Linked Glycans in Cancer	21
1.2.1.2.2. Core Fucosylation	23
1.2.1.2.3. Sialylation	26
1.3. Aberrant Glycosylation in Hepatocellular Carcinoma	30
1.3.1. Liver Cancer Glycomics	30
1.3.2. Liver Cancer Glycoproteomics	36
1.4. Mass Spectrometry	39
1.4.1. General Principles of Mass Spectrometry	40
1.4.1.1. MALDI Imaging Mass Spectrometry	41
1.4.1.2. Liquid Chromatography Coupled Tandem Mass Spectrometry	43
1.4.2. Mass Spectrometers in Dissertation	45
1.4.2.1. Bruker Solarix 7T FTICR	45
1.4.2.2. Bruker RapifleX TissueTyper MALDI-TOF	47
1.4.2.3. ThermoFisher Orbitrap Fusion Lumos	48
1.5. Broad Overview	49
Chapter 2: Hypothesis	52
2.1. Specific Aim 1	54
2.2. Specific Aim 2	55

Chapter 3: Changes in Glycosylation of Hepatocellular Carcinoma via MALDI-IMS	57
3.1. Abstract	58
3.2. Introduction	59
3.3. Materials and Methods	61
3.3.1. Materials	61
3.3.2. Tissues and Tissue Microarrays	61
3.3.3. Washes for Deparaffinization	62
3.3.4. Enzymatic Digestion and Matrix Deposition	63
3.3.5. N-Glycan Imaging using MALDI-IMS	63
3.3.6. Lectin Histochemistry	64
3.3.7. Statistical Analysis	65
3.4. Results	66
3.4.1. Increased Complex and Fucosylated N-Glycans in Liver Tissues	66
3.4.2. Analysis in Human Liver Tissue Microarray Set	70
3.4.3. Association of Specific N-Glycans with Survival	82
3.5. Discussion	84
Chapter 4: Determination of Core versus Outer Arm Fucosylation via MALDI-IMS	90
4.1. Abstract	91
4.2. Introduction	92
4.3. Materials and Methods	94
4.3.1. Cloning, Expression, and Purification of Endo F3	94
4.3.2. In-Solution Digestion by Endo F3	95
4.3.3. Glycan Sequencing	95
4.3.4. On-Slide Tissue Preparation and Imaging	96
4.3.5. N-Glycan Removal	97
4.4. Results	98
4.4.1. In-solution Analysis of Endo F3 Activity on N-Glycans	98
4.4.2. On-Tissue Analysis of Endo F3 Digestion via MALDI-IMS	100
4.4.3. Endo F3 Application to Patient Tumor Microarray	110
4.4.4. Core Fucosylation and Patient Survival Outcomes	113
4.4.5. Combined Application of PNGase F and Endo F3	115
4.5. Discussion	115
Chapter 5: Determination of Sialic Acid Linkage via MALDI-IMS	119
5.1. Introduction	120
5.2. Materials and Methods	122
5.2.1. Tissue Samples and Relevant Materials	122
5.2.2. Washes and Rehydration	122
5.2.3. Ethyl-Esterification Chemical Derivatization	123
5.2.3.1. Reaction Schematic	123

5.2.4. Amidation-Amidation Chemical Derivatization	124
5.2.4.1. Reaction Scheme 1	127
5.2.4.2. Reaction Scheme 2	127
5.2.5. Tissue Preparation and N-Glycan Imaging	127
5.3. Results	131
5.3.1. Ethyl Esterification On-tissue	131
5.3.2. Amidation-Amidation On-tissue	135
5.3.3. Amidation-Amidation on FFPE Human Liver Tissue	138
5.4. Discussion	143
 Chapter 6: Enhanced Glycoproteomic Analysis of Liver Cancer Tissues	 146
6.1. Introduction	147
6.2. Materials and Methods	149
6.2.1. Tissue Samples and Relevant Materials	149
6.2.2. Imaging Characterized N-Glycans	150
6.2.3. Segmentation Analysis	150
6.2.4. Tissue Extraction and Enrichment	153
6.2.5. Tryptic Digestion	153
6.2.6. Protein Quantification	154
6.2.7. Sample Clean-Up	155
6.2.8. Orbitrap Fusion Lumos Analysis	156
6.2.9. Protein and Peptide Identification	157
6.3. Results	158
6.4. Discussion	163
 Chapter 7: Conclusions, Limitations, and Future Studies	 166
7.1. Overall Findings	167
7.2. Changes in N-Glycans of Hepatocellular Carcinoma via MALDI-IMS	167
7.2.1. Conclusions	167
7.2.2. Limitations and Future Research	169
7.3. Enzymatic and Chemical Characterization of N-Glycans for MALDI-IMS	171
7.3.1. Conclusions	171
7.3.2. Limitations and Future Research	173
7.4. Enhanced Glycoproteomic Analysis of Hepatocellular Carcinoma Tissues	175
7.4.1. Conclusions	175
7.4.2. Limitations and Future Research	176
7.5. Conclusions and Final Thoughts	177
 REFERENCES	 178

## ABSTRACT

CONNOR ALLEN WEST. Determination of N-Linked Glycosylation Changes in Hepatocellular Carcinoma and the Associated Glycoproteins for Enhanced Biomarker Discovery and Therapeutic Targets (Under the direction of RICHARD DRAKE)

With hepatocellular carcinoma (HCC) remaining as the fifth most common cancer in the world, causing more than 700,000 deaths annually, the need for reliable, early stage diagnoses and preventive treatments is crucial. While serum glycoproteins are hepatic in origin, making them excellent targets for HCC biomarkers, they can originate from both cancerous and non-cancerous regions and direct analysis of cancerous tissue itself is lacking. To counteract this, I hypothesized that direct tissue analysis combined with proteomic analysis could be utilized to identify more potential targets specific to HCC for early detection. This was done with a primary focus on glycosylation—as most clinically approved biomarkers are glycoproteins—and examined direct tissue glycomics in conjunction with glycoproteomic techniques through two specific aims: 1) Determining patterns of N-linked glycan changes in HCC tissue using MALDI imaging mass spectrometry to compare to previously published serum changes and 2) identifying glycopeptides containing changes in observed patterns of N-linked glycans in HCC samples using a targeted glycoproteomic approach. In Aim 1, HCC tissue was examined using MALDI imaging mass spectrometry to

verify changes in glycosylation via direct tissue analysis. Here, it was found that increased branching and fucosylation were directly associated with the cancerous tissue when compared to normal or cirrhotic. To further identify changes in glycosylation, two methods (one novel and one adapted for imaging) were implemented on tissue to further classify N-linked glycan isoforms through linkage analysis, specifically for sialic acids and core fucose. Again, it was shown that core fucose is most directly related to HCC tissue, thus confirming serum findings in the literature. For Aim 2, the novel method of determining core fucosylation was used in conjunction with glycoproteomic techniques to further elucidate the core fucosylated glycoproteins of interest. With the tag left behind following the enzymatic cleavage, targeted glycoproteomics was used to determine glycoproteins of interest while eliminating some biases inherent in the method, such as low ionization efficiencies for more complex N-glycans. This work outlines the first in-depth analysis of HCC tissue specifically regarding N-glycan changes, a novel application to determine N-glycan isoforms, and the application of these methods for glycoproteomic enhancement. With these findings, new trends in glycosylation related to the disease state could be further uncovered, as well as provide new biomarker candidates or therapeutic targets for future studies.

## LIST OF TABLES

Table 1: Key Studies Highlighting Changes in Glycosylation Related to HCC.	31
Table 2: Patient Characteristics for TMA #1.	71
Table 3: Patient Characteristics for TMA #2.	72
Table 4: Master List of N-Linked Glycans.	74
Table 5: Glycans Altered in HCC Versus Cirrhotic or Adjacent Tissue.	77
Table 6: Master List of N-Linked Glycans with PNGase F.	107
Table 7: Master List of N-Linked Glycans with Endo F3.	108
Table 8: Main Core Fucosylated N-Glycans Found in Prostate Tissue.	109
Table 9: Main Sialylated N-Glycans and Mass Shift from AA Reaction.	136
Table 10: Protein Identifications Containing Core Fucose Modification.	160
Table 11: Primary Biological Component or Pathway Containing Core Fucose Modification.	162



## LIST OF FIGURES

Figure 1: Hepatocellular Carcinoma Disease Staging and Diagnostic Options.	5
Figure 2: Monosaccharide Characteristics and Naming Systems.	15
Figure 3: Trimming and Processing of N-Linked Glycans.	18
Figure 4: Fucosylation Linkages and Enzymatic Activity.	25
Figure 5: Sialylation Linkages and Enzymatic Activity.	27
Figure 6: Sialyl-Lewis Structures and Associated Linkages.	29
Figure 7: Serum Analysis of Liver Tissue and Cells.	35
Figure 8: MALDI-Glycan Imaging of HCC Tissue.	37
Figure 9: Simplified Imaging Mass Spectrometry Workflow.	44
Figure 10: Simplified LC-MS/MS Workflow.	46
Figure 11: Workflow of Tissue-based Glycan Analysis.	67
Figure 12: Detection of N-Glycans in Normal, Cirrhotic, and HCC tissues.	68
Figure 13: Hematoxylin and Eosin Staining of Varying Tissue Types.	69
Figure 14: Representative Imaging Data from both TMA Datasets.	73
Figure 15: Total Patient Glycan Upregulation.	76
Figure 16: Analysis of Human Liver TMA #1.	79
Figure 17: Analysis of Human Liver TMA #2.	80
Figure 18: Lectin Staining Compared to MALDI-IMS Data.	81
Figure 19: Patients Demonstrating Elevated Levels of Fucosylated Glycans.	83
Figure 20: Survival Plots for Branched and Fucosylated Glycans.	85
Figure 21: SDS-PAGE Analysis of N-Glycans by PNGase F or Endo F3.	99
Figure 22: HPLC Analysis of Fetuin A N-Glycans Following Digestions.	101
Figure 23: Generalized Workflow of Endo F3 and PNGase F Treatments.	102
Figure 24: Full Mass Spectra for Prostate Cancer Tissues Treated with Endo F3 and PNGase F.	104
Figure 25: Prostate Cancer Tissues Analyzed with Multiple Enzymes.	105
Figure 26: Hematoxylin and Eosin Stain of Prostate Cancer Tissue.	106
Figure 27: Multiple Tissue Types Treated with Endo F3.	111
Figure 28: Patient TMA Treated with Multiple Enzymatic Digestions.	112
Figure 29: Survival Plots for Patient TMA using Endo F3 and PNGase F.	114
Figure 30: Mixture of PNGase F and Endo F3.	116
Figure 31: Ethyl Esterification Derivatization of Sialic Acids.	125
Figure 32: Reaction Schematic for Amidation-Amidation Reaction #1.	128
Figure 33: Reaction Schematic for Amidation-Amidation Reaction #2.	129
Figure 34: Stabilization of Sialic Acids by On-Tissue Ethyl Esterification.	132
Figure 35: Stabilization of Multi-Sialylated Tissue Glycans by EE.	134
Figure 36: Amidation-Amidation Chemical Derivatization On-Tissue.	137
Figure 37: Stabilization of Multi-Sialylated Tissue Glycans by AA.	139
Figure 38: AA Chemical Derivatization of HCC Tissue.	140
Figure 39: AA Chemical Derivatization of Human Liver TMA.	142
Figure 40: Example Segmentation Analysis.	151
Figure 41: Segmentation of Endo F3 Applied HCC Tissue.	152
Figure 42: Annotation of Representative MS/MS Spectrum.	159

## ACKNOWLEDGEMENTS

I would like to express my most sincere gratitude to all the people in my life that aided in my achievements throughout my time at the Medical University of South Carolina. First and foremost, I would like to thank Dr. Richard Drake for his guidance, intellect, and persistent humor that not only assisted in my growth as an independent scientist and researcher, but made it an enjoyable experience along the way. His dedication to ensuring a well-rounded graduate education through conferences, seminars, and instrumental trainings were an integral part of my development. A huge thank you to Drs. Mehta and Angel for their constant support, advice, wisdom, and tireless efforts in mentorship, behaving as my two “unofficial” co-mentors throughout the entire process. I would also like to thank the rest of my committee, Dr. Woster, Dr. Muise-Helmericks, and Dr. Olsen, for their guidance and assistance on my projects throughout my research career. A special thank you to all the members of the laboratory group, past (Kacey, Fred, Aaron, Harmin, Cameron, and Savanna) and present (Colin, Calvin, Kim, Grace, Vivian, Andrew, Sharon, Mengjun, Hongyan, Stef, Janet, and Danielle), for their support and willingness to help me succeed, and to the Ball laboratory and the Proteomics Core as well for their expertise in all things proteomics.

Along with my support professionally, my achievements would not have been possible without the support of my family and friends along the way. Thank you to my parents who believed in me and always showered me with encouragement despite never quite understanding what I was doing. Thank you to my dear friends in Charleston: Brad, Pam, Colleen, Amy, Kate, Miguel, Casey,

Connor, and Kelsey, who kept me sane and helped me achieve that precious work life balance so important to every graduate student. A special thank you to my best friend Sarah for her constant encouragement and support and her willingness to meet me whenever and wherever (quite literally) to ensure that I'm doing alright. A huge thank you to my "work wife" Alyson, because without her constant (and somewhat overbearing...) optimism, intellectual conversation, and outpouring of encouragement, I would have never made it through this program as successfully and happily as I did. Thank you to the best emotional support animals and "co-workers" during the quarantine, Melisandre and Tyrion. Your endless shenanigans never fail to make me smile and I can't imagine my life without your snuggles. Finally, the most sincere and genuine thank you to my fiancée Lia. There are no words to fully express my gratitude for all the venting sessions, constant bargaining for edits, late-night trips to lab when things break in the middle of the night, and general support and encouragement she has provided throughout the whole process. I couldn't have done it without her and I'm truly grateful to share these experiences with her by my side.

# **Chapter 1: Literature Review**

## **1.1. Introduction to Hepatocellular Carcinoma**

### **1.1.1. Background, Mortality and Standard of Care**

Liver cancer causes more than 700,000 deaths annually, making it the fifth most common cancer overall and second most common cause of cancer-related death worldwide [1, 2]. Over the last 10 years in the United States of America (USA), liver cancer has seen the greatest increase in mortality among any cancer type. In the Annual Report to the Nation on the Status of Cancer, between 1975 and 2012, mortality from liver cancer increased at an annual rate of 2.8 percent in men and 2.2 percent in women [3]. Indeed, the occurrence of liver cancer is predicted to continue rising in the United States and will exceed 50,000 cases by the year 2021. This will result in greater mortality rates than breast or colorectal cancer [4].

Hepatocellular carcinoma (HCC) is the most abundant form of liver cancer, occurring in 75 percent of all liver cases, followed by Cholangiocarcinoma in the bile duct (10 to 20 percent), and Liver Angiosarcoma (approximately 1 percent), and Hepatoblastomas [5]; 80 to 90 percent of HCC cases stem from cirrhotic livers, while the remaining 10 to 20 percent become cancerous directly from years of chronic inflammation or fibrosis. Men are three times more likely to develop HCC as women with cirrhosis-inducing risks higher in the male population, although chemically-induced HCC is still lower in women with the mechanism still not clearly understood [6]. Asian populations are most prone to development of the disease due to higher incidences of chronic hepatitis infections in the Asia-Pacific region and differences in viral infection age, followed

by African Americans, Hispanics, and then whites [7, 8]. Survival rates of HCC are abysmally low: Early-stage diagnoses have five-year survival rates of approximately 30 percent, while late-stage diagnoses are as low as 3.7 percent [9].

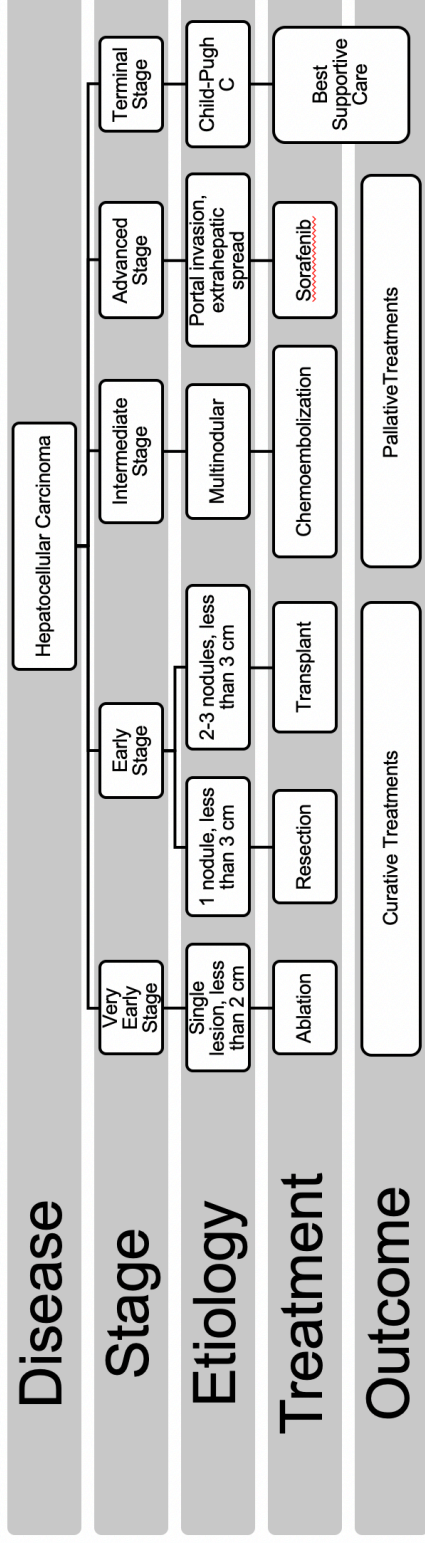
Related to the low survival rates of liver cancer patients, there are few treatments and even fewer curative options, especially for those patients with large lesions. While there are chemotherapeutic possibilities and ablation and resection techniques for lesions smaller than 3-5cm in size, there few curative options (survival longer than 60 months). The curative options currently available are surgical resection of small lesions and transplantation [10]. For each patient, a variety of factors such as hepatic reserve, hepatic function, and lesion size, determines which method is most viable. By most standards, patients with fewer and smaller lesions, as well as ample hepatic reserves, are often good candidates for resection, with five-year survival and disease-free survival rates at 39 and 26 percent respectively [10, 11]. Resection, however, is usually available to only 10 to 37 percent of patients at the time of diagnosis [11], and transplantation availability is even lower. Transplantation is the most successful form of curative therapy for liver cancer patients with overall and disease-free survival rates at 85 and 92 percent, respectively, but complications from immune rejection and a lack of organ donors results in transplantations being less common as a treatment technique [10, 11].

Chemotherapeutic options for liver cancer are limited and used primarily in those who are not candidates for resection. The frontline agent for those with

non-ablatable tumors is the multi-kinase inhibitor sorafenib, sold under the brand name Nexavar. Sorafenib is a general tyrosine and serine/threonine protein kinase inhibitor with activity against vascular endothelial growth factor (VEGF) and platelet-derived growth factor (PDGF) receptors as well as intracellular kinases B-Raf and Raf-1 [9]. Agents that specifically target one growth receptor, such as enhanced VEGF receptor inhibitors have failed to show activity against liver cancer [4]. It is noted that the activity of sorafenib against liver cancer is limited, with improved survival times of only a few months [12]. These bleak treatment options—both in their availability and efficacy—highlight the necessity for early detection of liver cancer, which allows for surgical intervention. Figure 1 details the clinically relevant outcomes of hepatocellular carcinoma as a function of time of diagnosis.

### **1.1.2. Liver Cancer Subtypes**

With the heterogeneity of hepatocellular carcinoma being as robust and diverse as it is, further classification was needed to accurately identify the mechanism of action for the disease. In 2009, a genomic-based study of approximately 600 western and eastern cases of viral-related HCC was done to develop a new class of molecular subtyping: S1, S2, and S3 [13]. Each classification correlated in terms of clinical parameters, but S1 showed dysregulation of the WNT pathway, which was surprisingly independent of  $\beta$ -catenin activation but rather dependent on the activation of TGF $\beta$ . S2 demonstrated increased proliferation and upregulated MYC and AKT activation.



**Figure 1: Hepatocellular Carcinoma Disease Staging and Diagnostic Options.** Representative diagram showing the different levels of staging in accordance with the Barcelona-Clinic Liver Cancer staging system, as well as the potential treatments and the effectiveness of these treatments



S3 was indicative of hepatocyte differentiation. Along with other factors, this classification system has been used to enhance biomarker targets through affected signaling pathways [14]. Through these different molecular classifications, therapeutic targets are more easily identified through the mechanism of which the disease operates; however, further studies are needed to elucidate truly viable targets for chemotherapeutic options.

Other groups have further expanded on this work, performing a comprehensive genomic analysis on hepatocellular carcinoma samples relative to the subtyping classification introduced by Hoshida [15]. This group performed clustering analysis and found three distinct clusters of molecular subtypes varying, differing from the previous subtyping classification system. The first integrated cluster found increased macrovascular invasion while containing the lowest fraction of differentiated samples based on the Hoshida classification, with relatively low amounts of CDKN2A silencing, CTNNB1 mutation, and TERT promoter mutation. These tumors, however, included more silencing of miR-122 and over-expression of MYBL2, PLK1 and MK167. Conversely, the second and third cluster experienced high CDKN2A silencing via DNA hypermethylation, higher rates of TERT promoter and CTNNB1 mutation. Cluster 2 contained lower grade tumors and less microvascular invasion while Cluster 3 had higher genetic instability and TP53 mutation. Cluster 1 was found to be associated predominately with subtype 2 and Cluster 3 was found to be primarily associated with subtype 3 for the Hoshida classification, further evaluating the underlying genetic modifications for the subtypes of HCC.

### **1.1.3. Disease Etiology**

#### **1.1.3.1. Risk Factors**

Being the most common form of liver cancer, HCC has many known risk factors, such as chronic infection with hepatitis B virus (HBV) and hepatitis C virus (HCV), alcohol abuse, obesity, aflatoxin exposure, and many other metabolic diseases [16, 17]. These risk factors induce a progressive inflammatory response, resulting in liver fibrosis and eventually cirrhosis, which is the true risk factor for HCC. This process occurs in multiple cycles of necrosis and regeneration, often leading to genetic instability [15]. Because of this genetic heterogeneity, the pathways involved in hepatocarcinogenesis are not fully clear, resulting in a lack of diagnostic and therapeutic options [18]. Therefore, the survival rates of primary liver cancer are low, generally with a 0.95 ratio of mortality to occurrence and five-year survival rates as low as 11 percent [19, 20].

Hepatitis viral infections remain the largest risk factor for liver cancer, with HBV accounting for 50 percent of all HCC and 70 to 80 percent of those cases arising from cirrhosis of the liver [21]. There is risk associated for those who test positive for the hepatitis B surface antigen (HBsAg), although the risk is still loosely correlated to the level of viral DNA. However, this risk can be compounded by other factors like family history, coinfection, and exposure to aflatoxins [22]. In contrast, HCV increases the chances of developing hepatocellular carcinoma 15- to 20-fold due to increased fibrosis and cirrhosis [23, 24].

Second to viral infection, environmental factors also play an important role

in the development of hepatocellular carcinoma. The most predominant risk, especially with the rise of the western diet, is that of obesity and diabetes which often lead to non-alcoholic fatty liver disease (NAFLD) [25]. Nearly 30 to 40 percent of men and 15 percent of women in the United States have some varying level of NAFLD, and this risk is increased with the presence of type 2 diabetes mellitus (T2DM) [26]. NAFLD also contains a more severe form, non-alcoholic steatohepatitis (NASH), which increases the risk of developing cirrhosis and liver related diseases.

Also, associated with HCC development is aflatoxin exposure, a natural mycotoxin produced by the *Aspergillus* species. Aflatoxin B1 (AFB1) is the most carcinogenic for HCC development, involved in the cytochrome P-450 system and forming AFB1-albumin, AFB1-guanine and other protein adducts responsible for carcinogenic hepatic DNA mutations [27].

### **1.1.3.2. Common Signaling Pathways**

Multi-omic studies tackle the topic of characterizing HCC, including analyses in genomics, proteomics, transcriptomics, glycoproteomics, glycomics, and metabolomics. While some studies have a broad focus and touch on multi-omic approaches, others focus primarily on one to further elucidate possible changes and therapeutic targets between the cancerous region, cirrhotic tissue, and normal tissue. In broad studies, the focus is often on building a network that links many aspects of the specific cancer to determine affected pathways. For example, Resson *et al.* (2016) characterized 499 genes, 217 proteins, 296

glycoproteins, 41 N-glycans, and 48 metabolites that represented significant changes between HCC and cirrhotic tissue, enabling the creation of a network that identified the most dysregulated pathways [28]. These findings demonstrated that tRNA charging, epithelial adherin junction remodeling, ILK signaling, EIF2 signaling, and glycolysis are significant pathways in the formation and maintenance of HCC. While these broad-scale studies usually don't lead to therapeutic targets, they provide a starting point for more specific -omic studies to move forward. Recently, one of the largest studies involved a genomic characterization of tissue across multiple platforms [15]. These researchers were able to corroborate and expand on the previous findings, determining genes altered more significant pathways such as  $\beta$ -catenin/WNT and RTK/RAS/PI(3)K and other factors such as TERT, TP53, CTNNB1 and immune checkpoints [15, 29, 30]. These genomic studies of tissue are increasingly important in HCC as knowledge of the tumor heterogeneity increases. It has been well studied that HCC displays frequent heterogeneous growth patterns and features, often within the same tumor, making it difficult to accurately determine a specific pathway or gene that fits precisely for each case [30-32]. With the successes of alpha-fetoprotein (AFP) as a viable serum biomarker, many studies have shifted to proteomic and glycomic studies of liver tissue in hopes of establishing a more encompassing method of detection or developing a therapeutic target.

#### **1.1.3.3. Origin of Disease Initiation**

As stated above, HCC frequently develops in the presence of cirrhosis,

and results in cellular dysregulation such as loss of cell cycle control, loss of senescence, and dysregulation of apoptosis [33]. Different risk factors result in HCC in a variety of ways, but generally, patients with one or more risk factors develop fibrosis of the liver, which leads to cirrhosis, and ultimately, liver cancer—most often HCC [33].

In the case of a viral infection, such as HBV or HCV, modes of initiation vary widely. For example, in the presence of HBV, many cell signaling pathways are affected by the virus such as a decrease in differentiation, an increase in proliferation, more genomic instability, and an increase in fibrogenic qualities [34]. These affected cellular pathways often lead to dysregulation of oncogenes and cell cycle regulators, resulting in increased chances for HCC development and complications. In contrast, HCV is a little more complicated when examining its role in HCC progression. The highly heterogeneous virus contains a multitude of subtypes, and while some have documented the increase rate of HCC in specific genotypic subtypes, the exact mechanism of HCV-induced HCC is still argued [35]. While it is generally agreed upon that HCV-induced HCC is the result of persistent inflammation and viral interference of cell signaling and must occur in the presence of cirrhosis, some have recently demonstrated a relationship between HCV-induced EGFR-ERK signaling and the progression of HCC [36, 37].

In the case of a non-viral risk factor, such as NAFLD, inflammatory responses are generally the culprit for disease progression, with inflammation through the NF- $\kappa$ B pathway being largely responsible for the approximately

twenty percent of patients that develop fibrosis, progressing to cirrhosis, and ultimately HCC [38, 39].

#### **1.1.4. Current Detection and Diagnosis of Liver Cancer**

##### **1.1.4.1. Clinical Detection Approaches**

Current guidelines by the American Association for the Study of Liver Disease (AASLD), National Comprehensive Cancer Network (NCCN), and Department of Veterans Affairs (VA) recommend HCC surveillance with abdominal ultrasound (US) with or without AFP every six months in all patients with cirrhosis [11]. Although there is no randomized trial evaluating HCC surveillance in patients with cirrhosis, several prospective cohort studies have demonstrated an association between HCC surveillance and improvement in early detection and survival in patients with cirrhosis, after adjusting for known confounders and lead-time bias [40, 41]. Although the surveillance has efficacy, the majority of the patients in the USA are diagnosed beyond the early stage when curative therapies are no longer effective. In addition to poor sensitivity for early HCC detection, US and AFP are both prone to false positive results, leading to unnecessary patient anxiety and diagnostic testing [42, 43]. While some providers use alternative, expensive imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) in all cirrhosis patients (despite a dearth of supporting data), others have abandoned HCC surveillance from frustration about the poor accuracy, leading to underuse of HCC screening in clinical practice [44]. US is increasing in traction given the lack of contrasting

agents needed for accurate detection of HCC, improving the diagnosis in the elderly population where contrasting agents are restricted [45]. This methodology, however, is still not as accurate as needed and still not cost-effective for many. Given the importance of early tumor detection for improving survival among HCC patients, there is a need for surveillance tests with higher sensitivity and specificity.

Biomarkers were put into practice to enhance earlier detection through less invasive means. As mentioned previously, alpha-fetoprotein (AFP) is a widely used and clinically approved biomarker for detection of HCC. While commonly associated with gestation, this glycoprotein is often monitored in pregnant women while the fetal liver produces AFP throughout gestation and achieves normal adult levels by 8 to 12 months [46]. The role of AFP in humans is not widely understood, as it does not bind estrogen as with other organisms, however various isoforms have been shown to be promising biomarkers in the field of HCC progression with changes in the glycosylation site increasing its power to indicate of a cancerous state [47-50]. Comparison of the L1 isoform, which is not related to HCC, to the L3 isoform containing the additional core fucose residue, which is associated with the malignancy of HCC, can be helpful in determining cancerous presence and degree of severity. AFP measurements can be taken directly from serum, allowing for a less invasive and more cost-effective screening method [51-54]. Combining AFP detection with US screening increased screening sensitivity to 90.2 percent, making this combination the most preferred method for the detection of HCC [55]. Recent reports have indicated

that algorithms consisting of several clinical factors and patient information can be used to improve the performance of AFP [56, 57]. It is also noted that AFP may be associated with very specific types of HCC [30].

As with many cancers, outcome is greatly improved by early detection [58, 59]. Overall survival of those detected with early cancers is less than 60 months but less than 20 months if the cancer is caught at a later stage [60]. Hence there is great significance in the development of methods for the early detection of HCC.

## **1.2. Glycosylation**

### **1.2.1. General Principles of Glycosylation**

Glycosylation, or the covalent addition of a carbohydrate chain to a protein, occurs through site-specific and enzyme-directed modification post- or co-translationally [61-63]. Glycosylation occurs in four major forms: N-linked, O-linked, glycosphingolipds (GSLs) and proteoglycans/glycosaminoglycans (GAGs) [64]. N-linked glycosylation involves attachment of the carbohydrate chain to an asparagine residue with a consensus sequence of N-X-S/T (where X can be any amino acid except proline) and O-linked glycosylation is the attachment of a carbohydrate chain to a serine or threonine residue. This modification occurring on cell surface proteins is crucial for cell-cell adhesion, signaling, and other cellular processes [65] and because of this dynamic variability, it is often a target for investigation as many disease states alter glycosylation expression [4, 52, 66]. Alteration of glycosylation can occur in many forms, such as overexpression









of specific glycoproteins associated with certain glycans, an increase or decrease of available sugar donors, and a change in glycosyltransferase and glycosidase enzymes [67]. Figure 2 details the different sugar residues involved in N- and O-linked glycosylation, as well as the different nomenclature used in describing glycans.

Aside from the forms of glycosylation listed above, there are other forms of glycoconjugates, such as proteoglycans and glycosphingolipids. Proteoglycans are built from a single or multiple glycosaminoglycans (a linear chain of repeating acidic disaccharide units) such as hyaluronan, chondroitin sulfate, keratin sulfate, and heparan sulfate [68]. Glycosphingolipids, on the other hand, are glycans attached to a lipid ceramide, which tend to contain a varying amount of core structures and gangliosides, and have been shown to regulate receptor tyrosine kinases [69]. In this dissertation, I will focus primarily on N-linked glycosylation and its physiological, pathological, and functional role in mammals, as well as its role specifically in HCC.

#### **1.2.1.1. Mechanism of Protein N-Glycosylation**

Protein glycosylation begins with the canonical hexosamine biosynthetic pathway (HBP) in the endoplasmic reticulum (ER). Along with playing a role in driving tumor growth and participating in the hallmarks of cancer, glucose, glutamine, fatty acids, and amino acids all play a role in the formation of uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) through HBP which acts as the basis for protein glycosylation [70]. Briefly described by Chiaradonna *et al.*, the

Monosaccharide	Symbol	Residue Mass	Oxford Nomenclature	Common Abbreviation	Generic Nomenclature
N-Acetylneuraminic acid		291.0954	S	Neu5Ac	Neu5Ac
Mannose		162.0528	M	Man	Hex
N-Acetylgalactosamine		203.0794	---	GalNAc	HexNAc
N-Acetylglucosamine		203.0794	A	GlcNAc	HexNAc
Fucose		146.0579	F	Fuc	dHex
Galactose		162.0528	G	Gal	Hex

**Figure 2: Monosaccharide Characteristics and Naming Systems.** Monosaccharide residues involved in glycan formation and synthesis are described above, as well as their masses, symbolic representation, and different styles of nomenclature

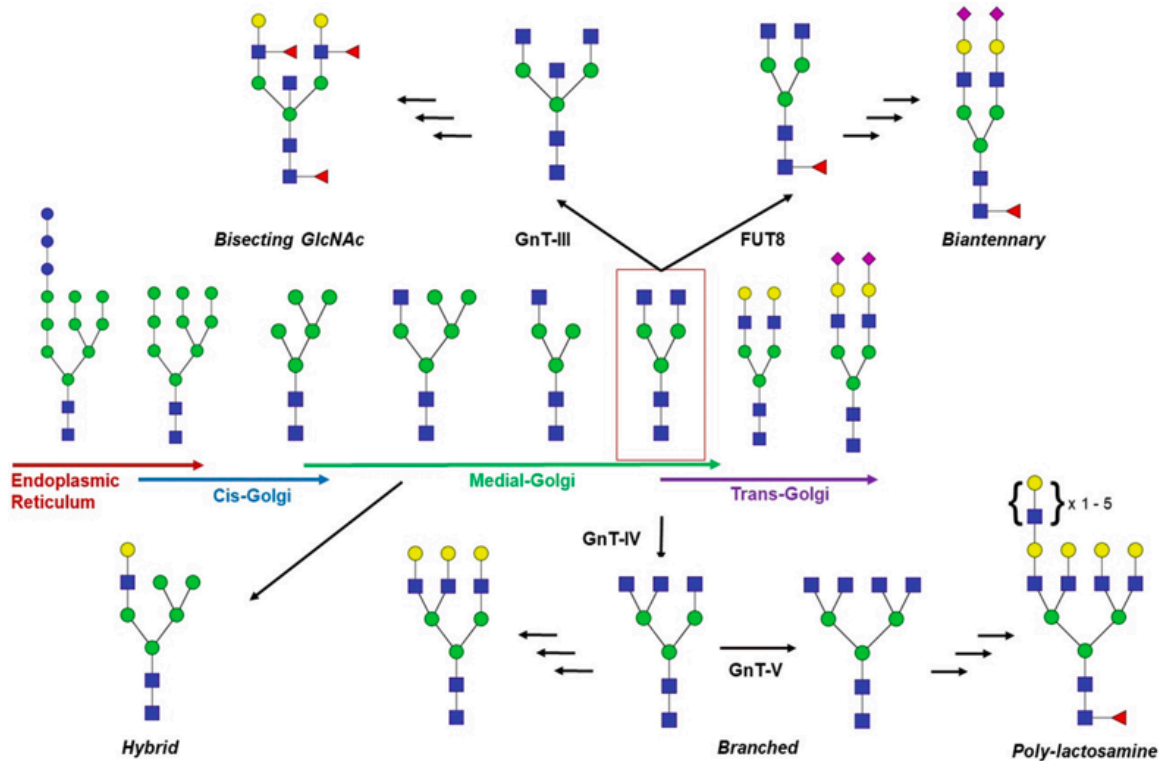
hexosamine pathway and glycolysis share the first two steps of glucose uptake and conversion to fructose-6-phosphate (F6P). Here, however, the paths diverge and glutamine fructose-6-phosphate amidotransferase (GFAT) is used in this rate-limiting step to transform the F6P to glucosamine-6-phosphate (G6P) with a byproduct of glutamate [71]. It is at this time that glucosamine entering the cell can also be converted using the GlcNAc kinase (GNK). Next, glucosamine-phosphate N-acetyltransferase (GNPNAT) catalyzes with acetyl CoA to create N-acetylglucosamine-6-phosphate (GlcNAc-6P) followed by a phosphomutase to create N-acetylglucosamine-1-phosphate (GlcNAc-1-P). Finally, uridine triphosphate (UTP) and GlcNAc-1-P produce UDP-GlcNAc through the UDP-N-acetylglucosamine pyrophosphorylase (UAP1/AGX1) [71, 72].

To begin the synthesis of N-linked glycans, however, the UDP-GlcNAc must be used with the dolichol phosphate (Dol-P) to react with the UDP-GlcNAc glycotransferase to generate the precursor N-acetylglucosamine dolichol pyrophosphate (GlcNAc-Dol-PP) [73]. Once the membrane-bound GlcNAc-Dol-PP is synthesized, sugar residues are attached on the cytosolic side in systemic fashion via multiple enzymes to create the first half of the precursor oligosaccharide which consists of two N-acetylglucosamines and five mannoses. This is then flipped from the cytosolic to luminal side of the ER via a flippase to initiate the final addition of sugar residues to create the final precursor oligosaccharide of two N-acetylglucosamines, nine mannose, and three glucose units [74]. The precursor oligosaccharide is then transferred from the Dol-PP onto an asparagine residue of a protein in the consensus sequence of N-X-S/T

where X cannot be proline.

Following the attachment of the precursor oligosaccharide, the N-linked glycan is then processed and trimmed. There are two different pathways by which this can be done: the glucosidase-independent and -dependent pathways. In the glucosidase-dependent pathway, processing occurs within the ER and is initiated by the trimming of glucose residues through the calnexin/calreticulin cycle. In contrast, the glucosidase-independent pathway occurs in the Golgi and utilizes an endomannosidase to cleave the Glucose1-3Mannose1 residue. The glycan is then processed to either create varying mannose structures or trimmed fully to five mannose structures and two N-acetylglucosamine structures. These glycans are then rebuilt to create more complex sugar structures of varying degrees and complexities [74, 75].

As shown in Figure 3, there are a multitude of pathways where glycan reassembly can occur with different complexities and additions available. As stated above, the first available route would be for the N-glycan to become a high mannose type glycan within the cis-Golgi. This is a glycan that contains a varying degree of mannoses, typically between five and nine, and stems from the removal of mannoses from the preliminary oligosaccharide. The mannose are trimmed within the cis-Golgi and if the N-glycan moves beyond a high mannose type and is fully trimmed, it is transferred out of the cis-Golgi and into the medial-Golgi. In the medial-Golgi, the N-glycan faces another turning point. At the first addition of an N-acetylglucosamine (GlcNAc), it will either continue to be processed to remove high mannose, or the GlcNAc will be built upon to create



**Figure 3: Trimming and Processing of N-Linked Glycans.** Schematic shows the trimming and processing of N-linked glycans following attachment to the protein, as well as the possible pathways N-linked glycan processing can follow

what is known as a hybrid-type glycan: a half mannose, half complex glycan. However, if the mannoses are fully removed, the N-glycan moves from the medial-Golgi to the trans-Golgi where the final addition of sugar residues occur. Here, there are multiple options for the glycan to take: fucosylation, more branching, bisecting, and sialylation. Fucosylation occurs with the addition of a fucose residue through one of the many fucosyltransferases; increased branching occurs through the action of N-acetylglucosaminyltransferase (GnT) IV or GnT-V to create extra arms of GlcNAc and galactoses; bisecting occurs through GnT-III and adds a GlcNAc to the first mannose of the structure, creating a bisecting branch of the glycan; and finally, sialylation occurs through the galactose sialyltransferase (STGal) family of adding a sialic acid onto a galactose, or the sialic acid sialyltransferase (STSia) family for addition onto another sialic acid [75-77].

Since this process is complex in nature, it is no surprise that many congenital disorders of glycosylation are likely to occur within this process, and many of these diseases arise from the dysregulation of enzymes responsible for glycosylation biosynthesis [74]. Some major classes of these congenital disorders of glycosylation (CDGs) within the N-linked glycosylation pathway are PMM2-CDG, characterized by dysregulation of the phosphomannomutase 2 gene; MPI-CDG, characterized by the dysregulation of the phosphomannose isomerase gene, and ALG6-CDG, dysregulation of the ALG6 gene resulting in aberrant attachment of the final three glucose molecules [78]. These diseases can vary in severity and effect, ranging from intellectual defects and

developmental delays of the individual to hormonal and chemical deficiencies throughout their life. Increased understanding of the role of glycosylation has become of particular interest to many fields, not only for the physiological roles, but also for the pathological role that glycosylation can play.

#### **1.2.1.2. Physiological and Pathological Roles of N-Glycosylation**

As stated above, glycosylation plays a vital role in many physiological pathways such as cell-cell adhesion, protein folding, transport, and cellular signaling [79]. In examining the major functions of N-linked glycosylation, the most relevant would be its major function: regulating protein folding. It is shown that when glycosylation is inhibited in eukaryotic cells, the most affected process is proper protein folding where proteins generated accumulate in a misfolded and nonfunctional state [80, 81]. For example, in-vitro studies have shown glycosylated and non-glycosylated versions of the same protein have different folding processes and N-glycans have been shown to alter conformational preferences near the site of glycosylation and move the protein towards more compact conformations, such as a compact  $\beta$  turn in the secondary structure [82]. Glycosylation has also been shown to act as a chaperone for incomplete proteins to enter the calnexin-calreticulin cycle to prevent movement from the ER to the Golgi apparatus until the protein either becomes fully folded or is degraded [74].

Another major physiological role of N-glycosylation is its involvement with protein transport and targeting. In the secretory pathway, glycosylation also

assists in deciding when and how quickly proteins are secreted, as well as signaling for other proteins involved in the secretory pathway [82].

Finally, the last major role that glycosylation plays is coordinating cell-cell interactions, both intrinsically and extrinsically. Glycans play a role in cell to cell recognition as well as cell-matrix interactions, as is the case for the adhesive selectin family, or the recognition of sialyl-Lewis X (sLe<sup>x</sup>) by the family of selectins [80]. Carbohydrate interactions can also occur, acting as an adhesive to maintain biological interactions. [83]

As is the case with most broadly involved biological processes, this leaves a lot of room for error in systemic function, especially in terms of glucose regulation. It is due to this link that glycosylation has been shown to play a major role in pathological functions, not only related to CDGs, but in terms of tumor progression, metastasis, and non-mediated cell proliferation and other mechanisms of disease and overall human health [84-86].

#### **1.2.1.2.1. N-Linked Glycans in Cancer**

For many years, glycosylation has been well characterized and functionally studied, but it wasn't until recent years that glycosylation has gained increased traction in the study of cancer, even though the first link between oncogenic transformation and glycosylation was described as early as 1949 [87]. Impaired glycosylation occurs in two principle mechanisms: incomplete synthesis and neo-synthesis. Incomplete synthesis typically occurs in earlier-tagged cancers and results in truncation of complex glycans due to dysregulated



glycosyltransferases, while neo-synthesis is more common in advanced cancers and involves the induction of genes related to glycan biosynthesis, resulting in the increased expression of certain glycans [68]. These alterations in glycosylation typically correspond directly with nearly all the hallmarks of cancer (established by Weinberg in 2000) such as the evasion of growth suppressors, the dysregulation of metabolism, resistance to apoptosis, immortality, increased invasion, and metastasis [67, 88].

Briefly examining the broader implications of glycosylation and its roles on each aspect of the hallmarks of cancer, trends emerge in the role glycosylation plays in cancer progression and maintenance. Beginning with proliferative signaling, studies have shown that N-glycan branching can mediate growth factor receptors to signal proliferative signaling [89] while numerous other growth factors are shown to be regulated through glycosylation as well [69]. Along with this, cell growth and survival has been shown to be regulated through glycosylation as well modification of the signaling for CD44 [90, 91]. In examining invasion and metastasis, dysregulation of MGAT5—responsible for the fourth branch of an N-glycan structure—has demonstrated a clear link for disruption of E-Cadherin, resulting in non-functional adherence junctions and cell-cell adhesion impairments, while MGAT3—responsible for the bisecting GlcNAc—has been shown to influence interactions with galectins and growth factors [92-96]. Finally, N-glycosylation has been demonstrated to be involved with the inflammatory pathway, with selectin binding specific glycosylated epitopes, initiating an immune response from leukocytes to the region of dysregulated

glycosylation [97, 98].

As mentioned, glycosylation plays many roles in cancer progression, evasion, and metastasis, but further studies have revealed that specific moieties of N-glycans are the most highly associated with the hallmarks of cancer. The first class of glycans are those with increased core fucosylation, shown to have high metastatic potential in a variety of cancers [99-101]. The second major class of glycan are those that have increased levels of Sialylation to disrupt cellular adhesion as well as increase cellular signaling for a variety of factors [102]. Below, these two major classes and their specific relation to HCC will be discussed at greater length.

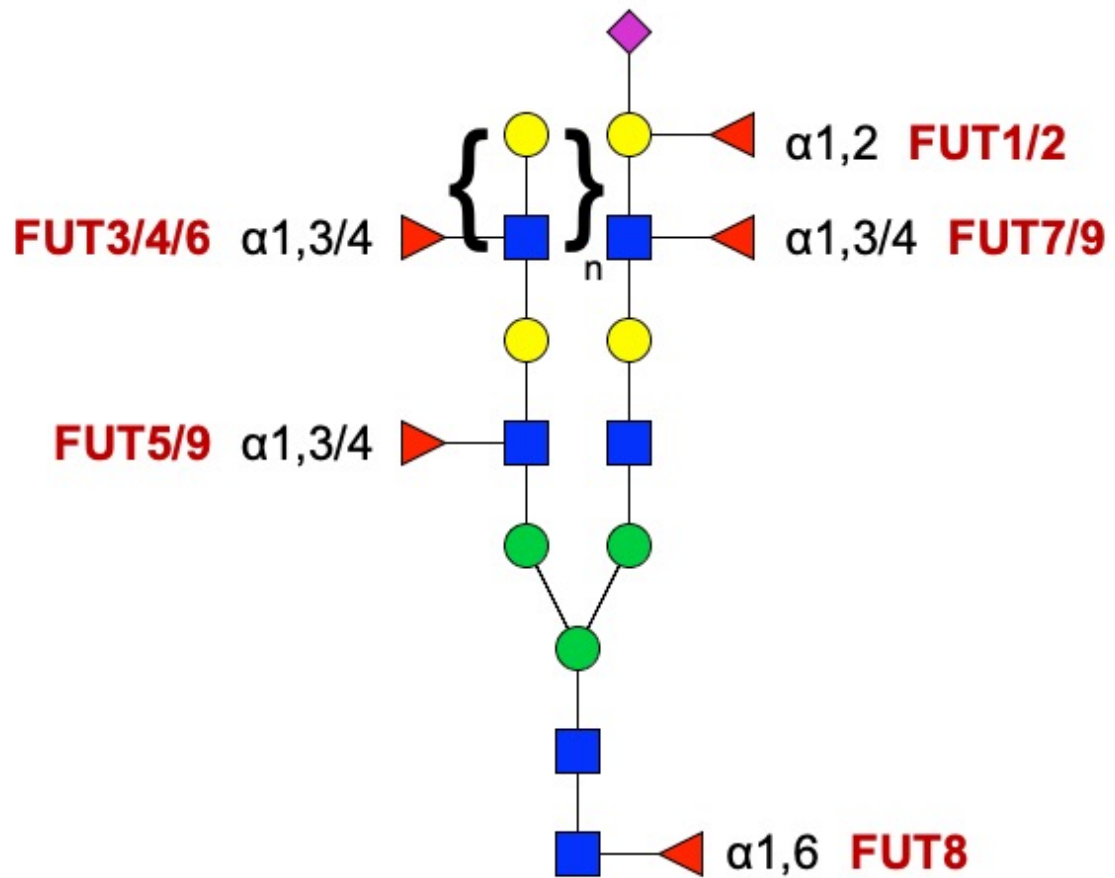
#### **1.2.1.2.2. Core Fucosylation**

The first major glycan class found to be associated with physiological and disease pathways, including HCC, is that of core fucosylation. Fucosylation is the addition of a fucosyl sugar moiety, which is transferred onto a glycan structure from a guanosine diphosphate (GDP)-fucose via enzymatic activity. Therefore, dysregulation of fucosylation can occur through either dysregulation of the enzyme responsible or through the available synthesized GDP-fucose [103]. There are two separate types of fucosylation that can occur on an N-glycan: core fucosylation, where the fucose sugar moiety is attached at the base of the N-glycan in an  $\alpha$ 1,6 linkage, and outer arm fucosylation, where the fucose sugar is placed anywhere else on the glycan at an  $\alpha$ 1,2/3/4 linkage. In humans, there are 11 fucosyltransferases responsible for the addition of fucose on an N-linked

glycan; however, there is only one enzyme, termed fucosyltransferase-8 (FUT8), that is responsible for the core fucosylation [104]. Figure 4 demonstrates the potential sites of fucosylation on a complex N-glycan, as well as the potential linkage conformations of the fucose sugar residue.

The role of core fucosylation has been vastly studied, with effects following directly in line with those seen in glycosylation overall, such as cell signaling and inflammatory response. For example, core fucosylation has been shown to mediate the signaling of B cell receptors to activate signaling required for pre-B cell recognition [105, 106] and is essential for a multitude of growth factor receptor functions [107].

Indeed, with the role that fucosylation plays within the signaling and immune response realm, this has become a popular area of research in terms of cancer treatment options. Specifically looking at core fucosylation, we have seen a multitude of studies demonstrating an increase in core fucosylation being associated with cancers, such as the metastatic potential of melanoma or non-small cell lung cancer [99, 100], as well as a marked increase in HCC cases as well [101], leading researchers to hypothesize that fucosylation inhibition could potentially be a source of cancer therapy. Recently, an anti-fucosylation drug, known as 2-fluorofucose (2FF), was developed and tested for its efficacy towards treating cancer progression [108]. Excitingly, in HepG2 cells, 2FF showed a striking decrease in core fucosylation, suppressing cell proliferation and integrin-mediated cell migration, demonstrating the importance of core fucosylation in HCC formation and a promising lead for N-glycan based cancer therapeutics [109].

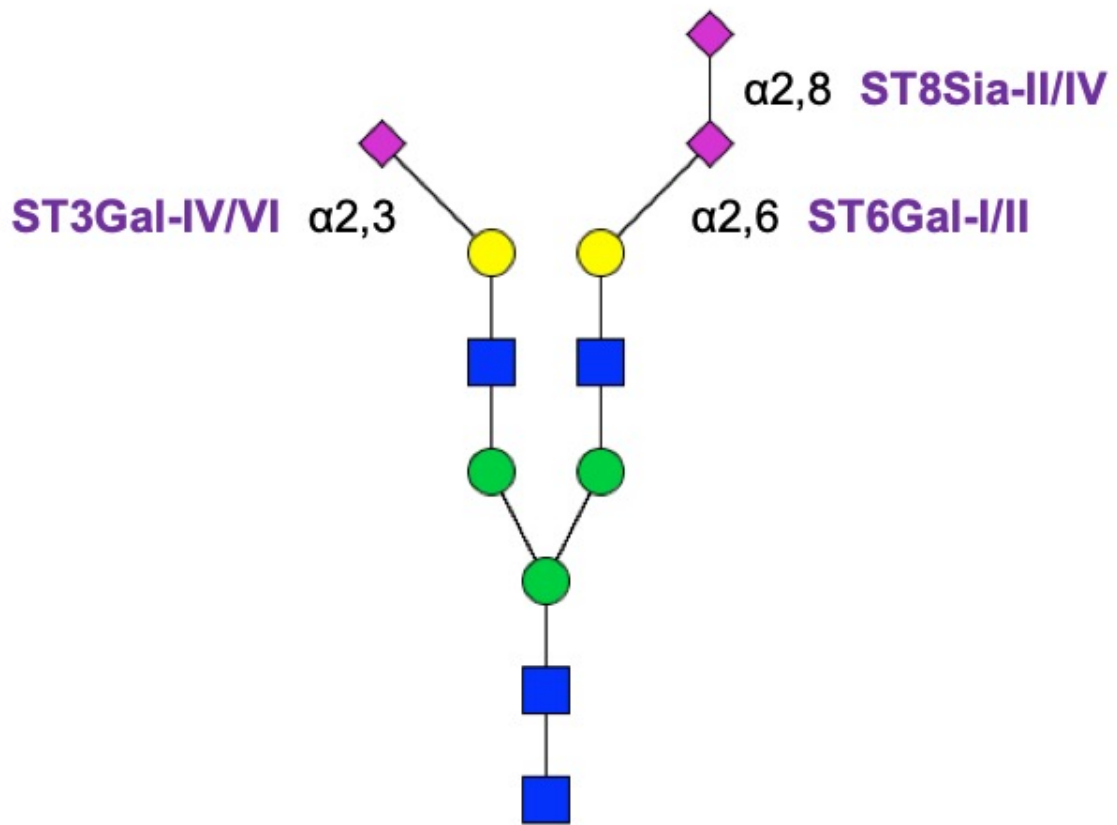


**Figure 4: Fucosylation Linkages and Enzymatic Activity.** Schematic shows the generic possible locations of fucosylation on an N-glycan, as well as the possible linkages in black and the possible enzymes responsible for this fucosylation in red. For clarity, FUT3/4/6 add a fucose to a non-terminal GlcNAc but not the first GlcNAc in a branch, while FUT7/9 add a fucose to the terminal GlcNAc in a branch

### 1.2.1.2.3. Sialylation

The second major class of glycan modifications is known as Sialylation. Sialylation is the addition of an N-acetylneuraminic acid (commonly referred to as Neu5Ac or sialic acid for short) onto either a galactose or another sialic acid via the action of a sialyltransferase (ST) acting on a CMP-Neu5Ac substrate within the Golgi. In humans, there are a number of STs that can be divided into two groups: those that catalyze the addition of sialic acid onto a galactose (ST6Gal-I, ST6Gal-II, ST3Gal-IV, and ST3Gal-VI) or onto another sialic acid to create a polySialylation chain on the glycan (ST8Sia-II and ST8Sia-IV) [77]. Similar to fucosylation, there are also additional linkage specificities associated with the attachment of sialic acid. Sialic acids attached to a galactose can be in either the  $\alpha$ 2,3 or  $\alpha$ 2,6 conformation, while sialic acids attached to another sialic acid will always be in the  $\alpha$ 2,8 conformation and must be attached to a sialic acid in the  $\alpha$ 2,6 conformation. Figure 5 demonstrates the potential sites of sialylation on a standard N-glycan, as well as the associated linkages and possible sialyltransferases for each potential site.

Similar to fucosylation, the role of sialic acids in human health has been well studied, as well as their role in varying disease states. Sialic acid is a unique monosaccharide as it is negatively charged at physiological pH, enabling different interactions and modifications. A primary example of the role of sialic acids in general human physiology is their relationship with lectin binding that control key processes in health in disease. Sialic acid binding lectins, known as siglecs, play a key role in modulating immune response pathways through toll-



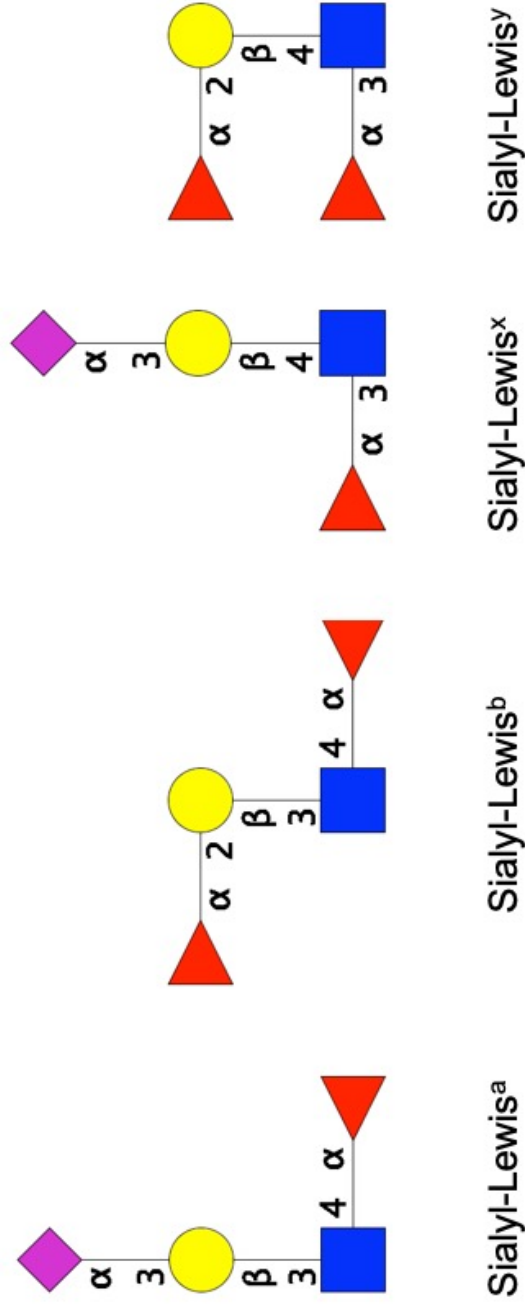
**Figure 5: Sialylation Linkages and Enzymatic Activity.** Schematic shows the generic possible locations of sialylation on an N-glycan, as well as the possible linkages in black and the possible enzymes responsible for this sialylation in purple.

like receptor signaling, CD33 regulation, and dendritic cell expression [110]. In the regulation of these signaling pathways, linkage of the sialic acid is of critical importance, emphasizing the need to better understand the specific linkages of these sialic acid modifications.

More interestingly, however, and relevant to current public interest, is the role sialic acids play in viral infection. For example, the human influenza virus is heavily reliant on the haemagglutinin trimer protein binding to  $\alpha$ 2,6-linked sialic acids to adhere to the host cell for infection [111] while in contrast, the betacoronaviruses utilize 9-O-acetyl-sialic acid as a receptor [112]. Therefore, sialic acid-based therapeutics have become increasingly prevalent, such as the invention of oseltamivir (Tamiflu) to act as a competitive inhibitor for cell surface sialic binding [113].

Aside from human health and viral infection, sialic acids have also been shown to play a role in cancer. Sialyl-Lewis X structures (Figure 6) are bound by selectins and mediate many functions of immune response and response to infection and injury. They become dysregulated in cancers, causing pro-inflammatory responses and leukocytic rolling arrests, allowing for cancer survival and metastasis [114]. Additionally, hypersialylation has been linked to increased metastasis and invasion of cancer, with upregulation of ST6Gal-I being linked to oncogenic Ras activation in cancers through altered  $\beta$ 1 integrin [115].

The role sialic acid modification plays in relation to HCC disease progression, metastasis, and immune invasion will be discussed later, as well as a closer look at the role of sialic acids in cancer.



**Figure 6: Sialyl-Lewis Structures and Associated Linkages.** Schematic shows the generic possible conformations for Sialyl-Lewis A, B, X, and Y, as well as the associated linkages of the monosaccharide components. Not pictured are the attached structure of the glycan, starting with the galactose attached to the GlcNAc.



### **1.3 Aberrant Glycosylation in Hepatocellular Carcinoma**

As stated above, many current biomarkers are glycoprotein biomarkers—such as AFP for HCC—with the glycosylation playing an important role in the detection of the disease. Previous work in serum by others has shown that many different structural motifs of these carbohydrate chains, or glycans, are associated with a disease state, such as increased branching, sialylation, fucosylation, or polylectosamine additions [52, 116-118]. Specifically for HCC, the addition of a core fucose ( $\alpha$ 1,6 linkage) to the associated N-glycosylation site on AFP is indicative of the disease [48]. Table 1 outlines some of the key studies related to the changes in glycosylation as it pertains to HCC.

Unfortunately, the use of these glycoprotein biomarkers is limited due to the lack of specificity for the tumor region. While serum is hepatic in origin and a viable option for biomarker detection of the disease, the sensitivity to detect early stage cases of HCC is still lacking. An approach to more effectively detect earlier cases of HCC with higher degrees of specificity and sensitivity could be more site-directed analysis of tumor and stroma directly in clinical tissue specimens.

#### **1.3.1. Liver Cancer Glycomics**

Glycomics has quickly become an emerging trend in the field of cancer biomarker development, and HCC is no exception [119-149]. In most cases, glycan analysis has been done with serum and not directly from the cancer tissue itself [122, 134-149].

Others have documented significant alterations in serum N-linked

**Table 1: Key Studies Highlighting Changes in Glycosylation Related to HCC**

<b>Glycan Modification Studies</b>	<b>Examples of Modification</b>	<b>References</b>
<b>Core Fucosylation</b>	AFP-L3, FUT8, GP73, miR122	Mehta, Herrera [4], Norton and Mehta [103], Zhou, Fukuda [109], Bernardi, Soffientini [150], Block, Comunale [151], Comunale, Lowman [152], Comunale, Rodemich-Betesh [153], Ma, Sanda [154], Wang, Sanda [155], Wang, Fukuda [156]
<b>Outer-arm Fucosylation</b>	$\alpha$ 1,3/4 fucosyltransferase, kinninogen, haptoglobin	Goldman, Resson [138], Comunale, Rodemich-Betesh [153], Mehta, Herrera [157], Kamada, Akita [158]
<b>Increased Branching</b>	GnT-V, GnT-IV, CD147/basigin-integrin $\beta$ 1 interaction	Kizuka and Taniguchi [159], Mehta, Norton [160], West, Wang [161], Cui, Huang [162]
<b>Sialylation</b>	MUC1, TF, sTF, Tn, sTn, ST6Gal-I, ST3Gal-IV, sialyl-Lewis <sup>x</sup>	Cao, Karsten [163], Chen, Wang [164], Gruszevska, Cylwik [165], Kongtawelert, Tangkijvanich [166], Mondal, Chatterjee [167], Powers, Holst [168]
<b>Polylactosamine</b>	JNK signaling, Stanniocalcin 1, CD147, $\beta$ 3GnT8	Kwan-Shuen Chan, Oi-Ning Leung [169], Liu, Qiu [170]

glycosylation with the development of HCC [171-176]. Specifically, the alterations most often observed are increased levels of alpha-1,3 and alpha-1,6 linked fucosylation found on bi-, tri-, and tetra-antennary glycans and, to a lesser extent, alterations in high mannose and polylactosamine glycans [122, 139, 143, 149, 151, 155, 173, 175, 177-191]. Many of these fucosylated proteins have been examined as biomarkers of HCC and are in the process of being commercialized. Results have shown that no one marker is sufficient to detect all HCC, but when these fucosylated glycoproteins are used as part of a diagnostic algorithm, area under the receiver operator characteristic curve (AUROCs) greater than 0.90 are obtained [155].

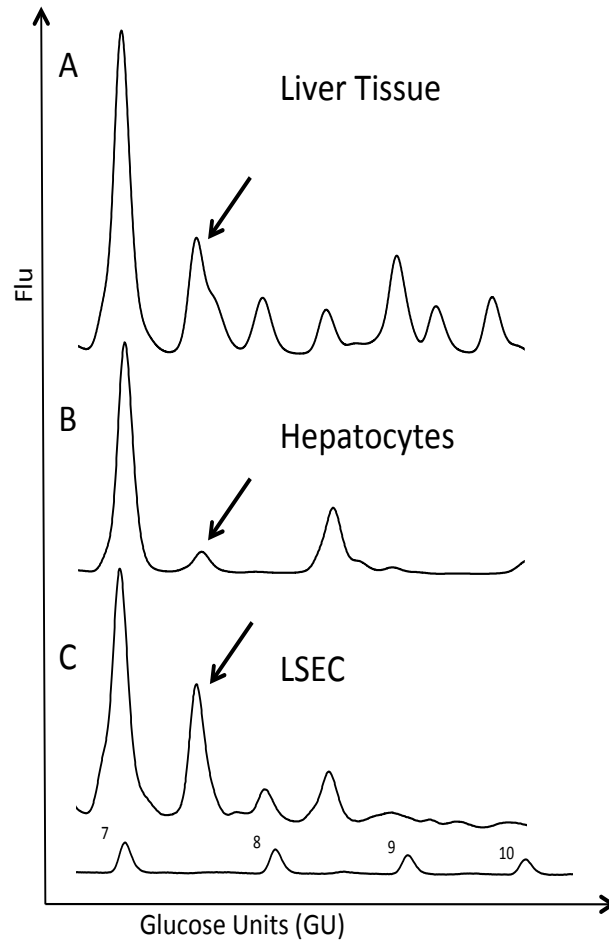
As stated, one of the most observed cancer-associated glycosylation modifications is core fucosylation, though the exact reason is still unknown. Enzymatic activity was one of the first possibilities explored, as the increase of FUT8 through the  $\beta$ -catenin/WNT pathway is seen in many cancers, including some HCC cases [192]. Tissue staining with core fucose-binding lectins has been used to determine the role of core fucosylation in HCC; however, when comparing normal, cirrhotic, and HCC tissue staining, there does not appear to be a significant increase of fucosylation solely within HCC tissue, but an overall trend within all tissue types [4, 193]. Along with core fucosylation, another glycan modification in HCC is increased glycan-branching, resulting in an increased presence of tetra-antennary glycans. These glycans are formed through  $\beta$ 1,6 N-acetylglucosaminyltransferase V (GnT-V or MGAT5) which results in an addition to tri-antennary glycans to form tetra-antennary structures [194]. This

modification, more so than core fucosylation, has been seen in HCC tissue specifically and could play a role in the cancer's development and metastatic potential [195-197].

Prior work about the source cells of serum fucosylation has been performed by several groups (including the Mehta Lab) with unclear conclusions. Initial work suggested that the genes involved in the production of GDP-fucose, the substrate for the fucosylation reaction, were increased in HCC as compared to surrounding normal tissue [198]. Support that transformed hepatocytes are the source cells for fucosylated proteins comes from the Mehta lab's recent work, where it was shown that as hepatocytes de-differentiate and undergo an epithelial to mesenchymal transition (EMT), they increase their level of fucosylation and up-regulate many of the genes involved in  $\alpha$ 1,6 linked fucosylation [199]. This is consistent with studies in lung cancer, where the  $\alpha$ 1,6-fucosyltransferase gene (FUT8) was involved in EMT [200]. It has also been shown that in a mouse model, deletion of FUT8 inhibits chemicals induced HCC by the down regulation of cancer-associated signaling pathways [156, 186]. Importantly, while these recent studies highlight the importance of fucosylation in cancer development, they do not offer any direct data on the source cell(s) for fucosylation in human disease.

Although a simple query, the fundamental question of the source of increased serum fucosylation has remained unanswered. As stated, Dr. Mehta's group has previously performed glycan analysis on HCC tissue following homogenization and HPLC-based glycan analysis [201]. In that study, two

surprising things were noticed: First, there was a much higher level of fucosylation observed in normal liver tissue as compared to human serum depleted of immunoglobulins (highly abundant non-liver derived serum protein). And second, while 8 out of 16 tissue pairs did have increased levels of fucosylation, statistically there was no change in fucosylated glycans when HCC tissue was compared to either normal liver tissue (from an independent liver) or from distal un-transformed tissue [201]. However, that study had two major flaws. First, there was no matching serum to allow for the analysis of both serum and tissue, so it could not be determined which of these patients had elevated fucose. Second, the glycan profile of hepatocytes and other liver cells have been determined and were found to be substantially different. That is, while liver tissue from normal individuals contains high levels of fucosylation (Figure 7A), purified human hepatocytes from the same individual have very little fucose (Figure 7B). In contrast, liver sinusoidal endothelial cells (LSEC)—another liver cell type—contain high levels of fucosylated glycan (Figure 7C). This high level of fucosylation within LSEC can confound the results when tissue is homogenized and examined in a mixed population. Therefore, although a HCC tumor may be primarily composed of transformed hepatocytes, adjacent liver tissue used for comparison will contain a mix of cells; thus, any comparison is not a “like for like” evaluation. This is true for glycan analysis, proteomic analysis, and expression data. This will be addressed for the first time using orthogonal methods. In regards to glycan analysis of tissue, recent laboratory members proposed a new method: MALDI imaging mass spectrometry (MALDI-IMS) [202]. Briefly, this



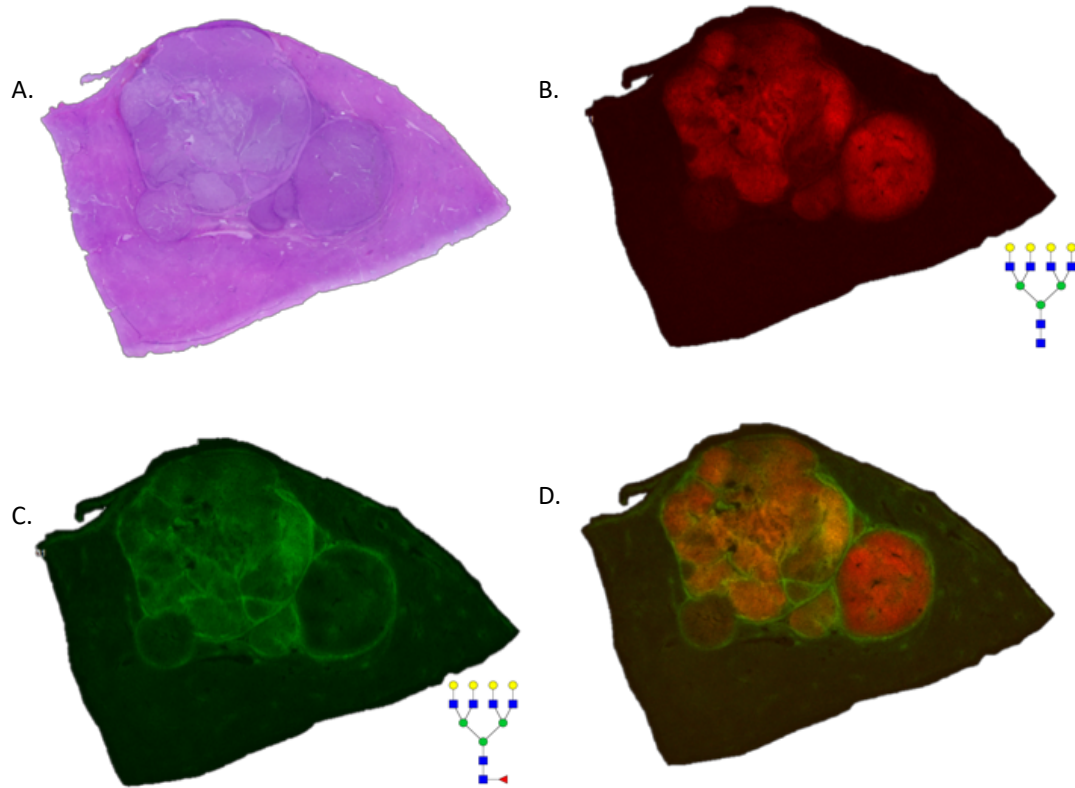
**Figure 7: Serum Analysis of Liver Tissue and Cells.** (A) Desialylated N-linked glycan profile of liver tissue from a normal individual; (B) purified hepatocytes from that same individual or (C) liver sinusoidal endothelial cells from that same individual (LSEC). Arrow points to the bi-antennary fucosylated peak, the only fucosylated peak observed in hepatocytes. While liver tissue has high levels of fucosylation, as do LSEC, hepatocytes have low levels.

method bypasses the need for microdissection and solubilization of tissue prior to analysis [160, 203-206].

Figure 8 presents an example of the type of data observed with MALDI-glycan imaging of HCC tissue. Figure 8A is the tissue following H&E staining and the large tumor is clearly visible, surrounded by non-malignant tissue. Figure 8B shows the distribution of one glycan, the tetra-antennary galactosylated branched glycan (A4G4), in red. As this panel shows, the A4G4 glycan is found predominantly in the tumor region with little observed outside of the tumor. Similarly, the fucosylated version of this glycan (A4G4F1), was also found predominantly within the tumor region (Figure 8C in green). However, as Figure 8D highlights, these glycans are differentially localized within the tumor, with the branched glycan without fucose predominantly in the inside of the tumor while the fucosylated branched glycan is on the outside of the tumor.

### **1.3.2. Liver Cancer Glycoproteomics**

Tissue proteomics have long been studied in many disease states, with excised tissue being homogenized and digested for protein analysis and comparative studies against normal tissue samples. Because of the availability of serum and its hepatic origin, many proteomic analyses have been done using serum, though tissue proteomics is of equal importance. Through these studies, links can be established between what is seen in serum and in tissue. Utilizing both top-down and bottom-up proteomics, researchers have been determining specific serum-derived proteins associated with disease state, tissue



**Figure 8: MALDI-Glycan imaging of HCC Tissue.** A) H&E stain of HCC tissue with surrounding non-malignant tissue. B) Localization of a glycan with a  $m/z$  value of 2393.840, which we have shown to be a tetra-antennary glycan. Red areas highlight localization of this glycan. C) Localization of a glycan with a  $m/z$  value of 2539.957, which has been shown to be a tetra-antennary glycan with a single fucose residue. Greens areas highlight areas of localization. D). Overlay of Panels B and C which show distinct localization of these glycans within tissue.



morphology, and progression.

As stated previously, AFP is a clinical biomarker that arose through proteomic research; however, others were also found to indicate HCC presence. Osteopontin, a biomarker that is measured in plasma, has also been studied in tissue and an increase in both osteopontin and Bcl-2 has been found in surgically resected HCC patients, indicating a co-dependence between the two in the tumorigenesis of HCC [207-209]. Along with osteopontin, peroxiredoxin 3 (PRX3) was also identified as a marker for HCC, and has shown an increase of expression on both the mRNA and protein levels in 94.9 percent of HCC cases [210]. In tissue analysis, PRX3 has been shown to indicate poor differentiation as associated with progression of the disease. Unfortunately, while these two markers have shown promising possibilities in the detection of HCC, they failed to detect HCC in the presence of high levels of cirrhosis, making them inappropriate for clinical application. With many proteins associated with HCC are found to be glycosylated, glycomic studies have become more relevant—both in serum and liver tissue analysis—to better understand the role and function glycosylation plays in HCC progression and the viability of glycoproteins as a therapeutic agent. More broadly in the glycoproteomic field, others have well characterized total serum glycosylation in HCC and found many associated modifications. For example, serum concentrations of core fucosylated haptoglobin have been examined to determine efficiency of detecting early stage HCC, as well as distinguishing HCC versus cirrhosis more efficiently than previously marketed AFP-L3 [211, 212].

Additionally, techniques to view glycoproteins and the associated glycan modification have become increasingly more prevalent, with targeted proteomics demonstrating its value in terms of serum evaluation and site-specific glycosylation. These studies have further confirmed changes in glycosylation for HCC as compared to normal, such as increased branching, hyper-sialylation and hyper-fucosylation causing an increased variety in N-glycosylation of haptoglobin in HCC and cirrhotic patients, or distinguishing those with chronic liver disease versus HCC [138, 213-216]. This will be further discussed later when examining the targeted glycoproteomics associated with cancer-specific HCC tissue.

#### **1.4. Mass Spectrometry**

Mass spectrometry has remained at the forefront of the analytical field for molecular analysis since its introduction in the late 1800s. Beginning with J.J. Thomson, he revolutionized the field in 1897 with the discovery of electrons and adopting the mass to charge ratio, where today these values are then converted to a mass spectrum that compares intensity of the ion to the mass to charge value. Mass spectrography was inducted into the scientific community in 1884 and used in Thomson's earlier experiments and in the early 1900s, more modern techniques were developed and the term "mass spectrometry" was coined to describe the field of study. Since then, the field has expanded to the variety of applications and techniques that we use today, such as drug testing for discovery and absorption properties; forensic analysis for trace amounts of evidence or explosive, arson, or drug abuse presence; environmental analysis to test water

quality, pollution contamination, and pesticide screening; and clinically, for drug therapy monitoring, peptide examination for diagnostic purposes, and many more.

#### **1.4.1. General Principles of Mass Spectrometry**

Mass spectrometry functions under the basic principle of conversion of the analyte of interest into gaseous ions that can then be sorted by their mass to charge ratios and relative quantities [217]. The intricacies of a mass spectrometer vary greatly depending on the type of mass spectrometer (more details below), but the basic components remain relatively standard: the ion source, the mass analyzer, and the mass detector system.

The ion source is involved in the ionization of the analyte and the transport of the ions to the mass analyzer. Simply, the ion source is the beginning of the mass spectrometry process where the ions are first generated [218]. Ionization is one of the key differences in techniques of mass spectrometry with a variety of options and types relevant to the instrument. Some examples of different ionization techniques used with biological samples are electron ionization (EI)—termed a hard ionization technique due to the fragmentation of the ion of interest, electrospray ionization (ESI), matrix-assisted laser desorption/ionization (MALDI), and desorption electrospray ionization (DESI)—all termed soft ionization techniques due to the little amounts of fragmentation and the residual amounts of energy imparted into the molecule of interest [219].

Once ions are generated, before they are detected and recorded, they will

pass through the mass analyzer, used to determine the characteristics of the ion and its mass to charge ratio ( $m/z$ ), a dimensionless parameter relating the elementary charges to the mass of the ion, and separate them according to these parameters using a generated field. Mass analyzers generally vary by the type of field generated (static or dynamic, as well as magnetic or electrical), and these factors are what account for the mass spectrometers resolving power, or how accurately it can separate the closest related ions, on different instruments [220].

Finally, the last component of mass spectrometry is utilization of the mass detector. The mass detector records the relative abundance of the ions being resolved and converts the charge or current produced into a signal, which is then converted into a mass spectrum to be analyzed. Mass detectors vary by instrument and amplification techniques are often required. The signal output can also vary based on the mass analyzer.

Below, I will outline a more in-depth analysis of the types of mass spectrometry pertinent to this dissertation, as well as details surrounding the specific mass spectrometers used.

#### **1.4.1.1. MALDI Imaging Mass Spectrometry**

The first major type of mass spectrometry used in this dissertation is MALDI imaging mass spectrometry (MALDI-IMS). In recent years, imaging mass spectrometry has emerged as one of the top areas of mass spectrometry and has nearly 900 related publications to date [221].

MALDI-IMS works under the basic principles of MALDI mass

spectrometry, but on a larger scale of adding a two-dimensional aspect to a sample, allowing for multiple points of analysis. Briefly, a Nd:YAG laser is used to strike the molecules coated in an ionizing matrix, resulting in a gaseous form of the analyte without fragmentation or decomposition. Most imaging mass spectrometers can analyze a variety of analytes, such as glycans, metabolites, drugs, lipids, proteins, and peptides, with the most important variable being matrix choice. Matrix choice varies depending on the need for the analyte, such as the ease of sublimation from the solid phase to the gaseous phase, or the absorption of the laser wavelength used by the instrument [222]. Examples of some matrix options used for imaging experiments, specifically in examining N-glycans, are 2,5-dihydroxybenzoic acid (DHB) and  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) [76]. In this dissertation, CHCA was used primarily to examine N-glycans in the positive ion mode.

Once the matrix is applied to the sample, the laser is used to ionize the sample in a pixelated fashion. The distance between the pixels is determined by the user pre-analysis and a mass spectra will be generated for each spot of the laser. Following completion of the run, every mass spectra collected will then be compiled into one average mass spectra that shows relative abundance of all peaks found throughout the analysis. Software such as FlexImaging (Bruker), can then be used to create a false color image at each peak of interest, allowing for pixels of higher relative abundance to appear more intense on the image, while pixels of lower relative abundance appear less intense, similarly to a heat map [223]. Figure 9 demonstrates a simplified approach to imaging mass

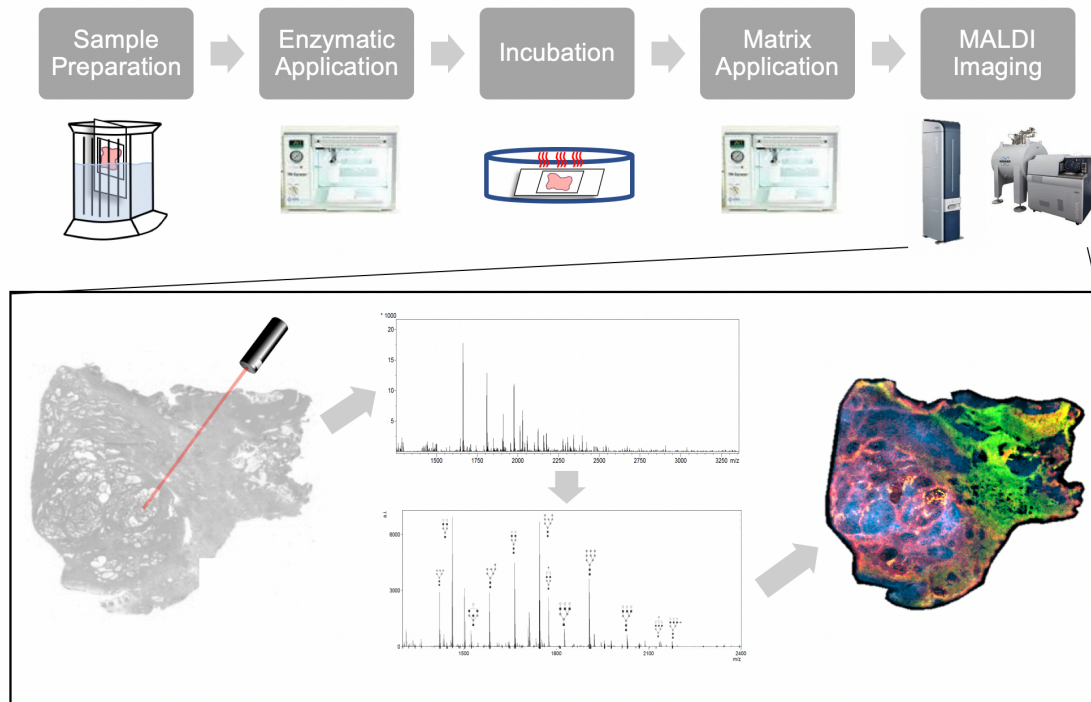
spectrometry and the data visualization process.

Imaging mass spectrometry offers several clinical benefits, combining the advantages of both mass spectrometry and microscopy together onto one platform. In general, processing times are lower than that of traditional immunohistochemistry and pathology processes and allow for multiple molecules to be analyzed at once, rather than a single analyte of interest. It is also more successful in analyzing histomorphologically ambiguous tissue regions, and able to discriminate classifications within clinical tissue more accurately. Improvements to the reproducibility of the method have grown with implementation of the mechanical sprayers, showing consistent reproducible data within complex biological replicates with errors reported in parts per million (ppm). With the clinical implications of this technique are continuously growing, the technique is approved through the Clinical Laboratory Improvement Amendments (CLIA) [224].

#### **1.4.1.2. Liquid Chromatography Coupled Tandem Mass Spectrometry**

In contrast to imaging mass spectrometry, liquid chromatography-mass spectrometry (LC-MS) offers a sensitive analysis of analytes and separation of these analytes. Synergistically combining mass spectrometry principles with the capabilities of a high/ultra-performance liquid chromatography (HPLC/UPLC) system allows for enhanced analytical capabilities.

The first portion of the analysis begins with the sample being separated



**Figure 9: Simplified Imaging Mass Spectrometry Workflow.** Diagram represents a simplified workflow for imaging mass spectrometry, specifically for an enzymatic imaging experiment. Top half represents the simplified steps in sample preparation to imaging, while the bottom half represents the actual imaging process and data analysis and results portion.

via a liquid chromatography step. Samples are injected into a mobile organic phase and utilizing a C18 stationary phase, peptides are captured and eluted on a gradient of the mobile phase, allowing for separation of analytes by mass. Following the separation of the analytes, the HPLC system is then coupled to a mass spectrometry system to undergo mass analysis of the analytes, similarly to what was described above in the general principles section. Typically, electrospray ionization is used to effectively couple the liquid phase analytes to the gaseous phase for mass analysis [225].

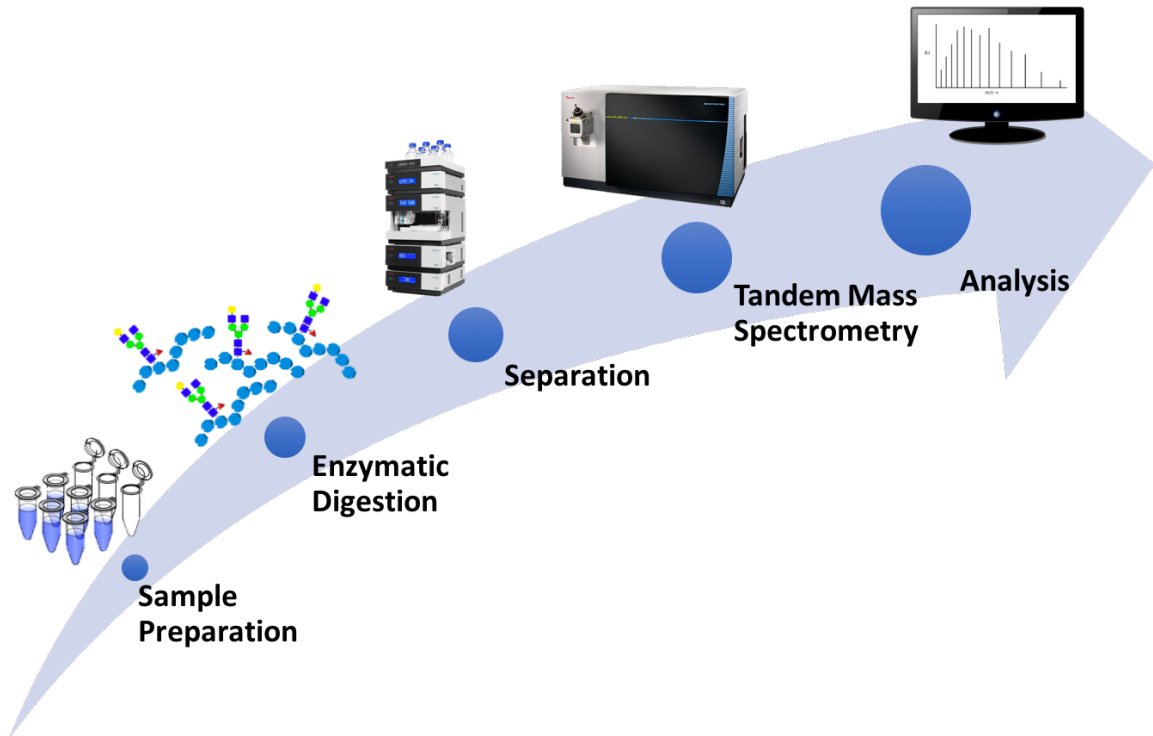
Often associated with LC-MS systems is the ability to do tandem mass spectrometry experiments (MS/MS or MS<sup>2</sup>). As stated above, the initial mass spectrometry analysis will separate ions by their mass to charge ratio; however, the second mass spectrometry analysis involves a further fragmentation of these ions for further quantification of analytes. This can be done through a variety of methods, such as collision-induced dissociation (CID), higher energy collision-induced dissociation (HCD), or electron-transfer dissociation (ETD), depending on the type of information desired [226-228]. Figure 10 shows a simplified model of the LC-MS workflow, as well as a more in-depth look at the mass spectrometer portion of an LC-MS/MS experiment.

## **1.4.2. Mass Spectrometers in Dissertation**

### **1.4.2.1. Bruker Solarix 7T FT-ICR**

The first mass spectrometer, and most frequently used in the dissertation,





**Figure 10: Simplified LC-MS/MS Workflow.** Diagram represents a simplified workflow for the LC-MS/MS protocol beginning with sample preparation, through enzymatic digestion and separation via HPLC, to the MS/MS experiment and data analysis.

is the Bruker Solarix 7T FT-ICR. Most of the mass spectrometry imaging analyses in this dissertation were performed on this instrument.

This mass spectrometer is unique in the fact that it is a MALDI Fourier Transform Ion Cyclotron Resonance (FT-ICR) instrument, allowing for higher mass resolution and resolving power, which comes with the cost of slower sample acquisition time. With the key metric being frequency, ions are better resolved, with the ability to separate more closely related ions, however this comes at the cost of a slightly longer analysis time. Briefly, ions are ionized and focused into a beam with DC and RF voltages, then proceed through the quadrupole, excluding ions outside of a specific  $m/z$  range of interest. Ions not excluded then pass to the hexapole. The ions are then focused into the ICR cell and the 7T magnet, where electrodes produce a cyclic motion of the ions to be converted to a frequency metric, causing them to separate, but keep a small electrical field to keep ions within the ICR cell. Finally, ions are excited by an RF frequency sweep and detected via the mass detector plate [229-231].

#### **1.4.2.2. Bruker RapifleX TissueTyper MALDI-TOF**

The second imaging mass spectrometer used in this dissertation is the Bruker RapifleX TissueTyper MALDI-TOF. This instrument comes equipped with improved laser technology for high throughput and better pixel-to-pixel reproducibility. This laser is capable of a 5 $\mu$ m focus, allowing for square imaging pixels with almost complete tissue coverage and no pixel-to-pixel interference. TOF mass analyzer is specialized for linear and reflector measurements in

positive/negative ion mode. This instrument was used after initial optimization for faster and higher spatial resolution imaging.

This instrument operates through a MALDI time-of-flight (TOF) system, allowing for faster acquisition time though a lower mass resolving power. Briefly, the laser induces an ionization particle cloud, known as the plume, and ions are extracted via a small electrical field. The free ions then pass into a field-free vacuum path until reaching the mass detector, where the time of flight is measured and used to calculate the mass to charge ratio, as typical matrices result in a standard charge state of one, causing smaller ions to separate from larger ions in a fixed field distance. The RapifleX offers two forms of mass detection: linear and reflector. In linear mode, ions travel down the path just once, but reflector mode reflects the ions back at a slightly different angle utilizing a reflectron. A reflectron creates a charged field for the ions to pass through, separating out similar mass clusters while also reversing the direction for the second part of the ion flight path. This allows for more ion separation and higher mass resolving power than a linear TOF instrument [232-234].

#### **1.4.2.3. ThermoFisher Orbitrap Fusion Lumos**

The final mass spectrometer used in this dissertation is the ThermoFisher Orbitrap Fusion Lumos LC-MS/MS system. This instrument is a tribrid mass spectrometer with HCD, ETD, and UVPD capabilities, as well as 1,000,000 full width at half maximum (FWHM) ultra-high resolution characterization. This instrument will be used primarily for glycoproteomic analysis of tryptic peptide

fragments still containing glycan structures attached. This will allow for peptide sequence, glycan structure, and site specific localization of the glycan on the peptide.

The first mass analyzer is the quadrupole (Q1) which allows for filtering of the ions, like described above. The second mass analyzer is the Orbitrap, which generates an electrostatic field and causes the ions to move in a complex spiral pattern. This current pattern is then Fourier transformed to create a mass spectrum, allowing the Orbitrap to function as a mass detector as well. Finally, the instrument contains a linear ion trap (LIT), which houses the ion storage, isolation, and collision-induced dissociation capabilities. The combination of the Orbitrap acting as a mass analyzer and detector, along with the pulsed ionization techniques that generate an extremely long flight path, results in a very high mass resolving power, while also introducing more advanced collisional fragmentation techniques, making this instrument ideally suited for proteomic analysis of biological material [218, 219, 235-237].

### **1.5. Broad Overview**

While tissue analysis of HCC has become more prevalent, there are still necessary steps required to link what is known regarding serum and tissue for more accurate biomarker discovery. Biomarker discovery and analysis is moving in the right direction, focusing more on the patient-specific tissue sections than simple circulating serum or plasma. However, most studies still fail to acknowledge the complex heterogeneity and morphology found within HCC

tissue. In many cases, a tissue block or section is obtained and homogenized for analysis. This method disregards all pathological and histological complexities within the tissue, often including normal adjacent tissue or cirrhotic tissue in the analysis. This is where many tissue analyses are lacking, in that they could possibly include patterns and expressions that are not associated with the cancer. This leads to challenges in the development of robust biomarkers, and contributes to their inability to detect earlier stages of liver cancer and disease. In the future, linking serum glycoproteomics to specific tissue glycomics within the cancerous region itself will become increasingly important. The utilization of MALDI imaging techniques provides a substantial basis for further analysis and helps determine protein and glycosylation changes in specific regions that can be correlated to disease states. In combining this technique with other –omic approaches, there is the possibility to develop more sensitive and specific biomarkers for enhanced detection of HCC.

The work outlined in this dissertation combines well-characterized and novel glycomic techniques in the field of imaging mass spectrometry and targeted glycoproteomic studies linked specifically to tissue histopathology. This novel work simultaneously addresses some challenges to the field in terms of tissue glycomics while maintaining all benefits afforded through this technique, such as determination of linkage specificity within an imaging mass spectrometry workflow, as well as further characterizing specific glycan modifications in cancerous tissue via targeted glycoproteomic techniques specific to areas of interest. Additionally, this work opens the door to many further investigations, as

this work was primarily studied in the liver and geared toward HCC, but these techniques could be applied to many other disease states and tissue types, enhancing all human glycome data and relationships to health and disease.

# **Chapter 2: Hypothesis**

While most constituents of serum are hepatic in origin and includes a clinically relevant biomarker for hepatocellular carcinoma (HCC), the sensitivity and specificity of biomarkers for early HCC cases are still lacking. One potential biomarker is core fucosylated alpha-fetoprotein (AFP-L3), which is FDA approved for the diagnosis of HCC, and there are now many other glycosylation variants that are being developed as biomarkers of HCC [153]. However, it is difficult to determine the exact source of these upregulated N-glycan modifications via circulating serum, giving rise to uncertainty of whether these modifications are originating solely from the cancer or other parts of the organ as well [238]. In the case of HCC, the heterogeneity of the cancer leads to variable expression for N-glycoproteoforms, contributing to the lack of early diagnostic methods, as current methods lack the required sensitivity and specificity for tissue specific biomarkers [239]. Therefore, it is imperative to find biomarkers more specific to the cancerous tissue itself. Matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry bridges this gap by providing a direct spatial analysis of the N-glycans in clinical samples. By linking currently studied glycomic modifications in cancers with spatially localized analysis of HCC tissue, it is possible to bridge the gap between overall tissue features and cancer specific trends. In doing so, it is also possible to determine associated glycoproteins for further diagnostic markers and therapeutic targets, shifting therapy from curative to preventative, especially before progression to late-stage cancer. The scope of this work provides a novel and informative look at site-specific N-glycan modifications within HCC. In general, the ability to associate known glycan



changes with the corresponding glycoprotein it was originally attached to is difficult due to sample complexity or the amount of sample needed. Here, by combining well-established MALDI imaging capabilities for spatial analysis on liver tissue with region-directed glycoproteomic analysis via MALDI imaging, enhanced biomarker capabilities are possible.

Here, I optimized a method for identifying more specific hepatocellular carcinoma biomarkers that combines previously studied serum glycoproteomics with novel methodology in tissue glycomics. Using this approach, we will be able to determine specific glycosylation changes in tissue and characterize glycoproteins associated with these changes in HCC tissue. Therefore, I hypothesized that glycomic tissue imaging using MALDI imaging mass spectrometry can be used in conjunction with glycoproteomic techniques to identify biomarkers for early detection and diagnosis specific to HCC. I tested this hypothesis through the following specific aims: 1) determining and characterizing the patterns of N-linked glycan changes in HCC via MALDI imaging mass spectrometry to compare to previously studied changes in serum and 2) using these identified changes in glycosylation to enhance targeted glycoproteomics for improved biomarker target identification.

## **2.1. Specific Aim 1**

*Determine patterns of N-linked glycan changes in hepatocellular carcinoma tissue using MALDI imaging mass spectrometry to compare to previously found changes in serum*

The overall goal of Aim 1 was to utilize the well-established MALDI Fourier transform ion cyclotron resonance (FT-ICR) and time-of-flight (TOF) imaging mass spectrometry method to examine N-glycosylation changes in specific HCC regions of liver tissue and validate these results with those previously studied in serum. In terms of MALDI imaging mass spectrometry, the ability to examine liver tissue in a spatially conserved manner is an innovation in the proteomic field [240]. Through enzymatic release of the N-glycans and matrix application, the N-glycans stay localized to the area they are released from and are free to be ionized and measured by mass spectrometry instrumentation [241]. This allows one to not only see the present structures and their location, but also give relative abundance and intensity of those glycans. As an added benefit, we can scan a whole range of mass values at a high sensitivity, providing more informative glycan data than other methods. Finally, localization of the analytes is used to determine pinpoint locations of interest within the HCC tissue for further glycoprotein analysis. Additionally, new enzymes and methods were further optimized in conjunction with the previously established method to characterize the N-glycans found in tissue, as well as differentiate mass based on N-glycan characteristic and structure.

## **2.2. Specific Aim 2**

*Identify glycopeptides containing changes in observed patterns of N-linked glycans in hepatocellular carcinoma samples via the ThermoFisher Orbitrap Fusion Lumos Tribrid Mass Spectrometer*

The overall goal of Aim 2 was to link what is known regarding the tissue glycomics via imaging and determine the glycoprotein and glycopeptides responsible for carrying the modifications of interest. For further glycoprotein analysis, a recently developed method utilizing Orbitrap mass spectrometry has been established. This method, known as HCD-PD-ETD (higher-energy collision dissociation-accurate mass- product-dependent electron transfer dissociation), allows for the examination of N-glycan structures and the associated protein it was originally attached to [227]. While this method is effective, it is inefficient for large scale studies, as the number of glycosylated proteins and associated glycans would be nearly impossible to analyze sensitively and with significance. However, by regionalized expression analysis via MALDI imaging, a more thorough and selective analysis can occur in only tissue regions of interest, generally those expressing the trends of N-glycan changes found specifically in HCC. By examining only these regions, the less-important proteins associated with HCC N-glycan modification were ignored and, therefore, reduced the possible suspects for potential biomarkers and therapeutic targets. This methodology increased our ability to perform glycoprotein studies, such as 1) looking at multiple proteins and glycoforms, 2) maintaining glycan information and associated protein information, and 3) confirming that the N-glycan or glycoprotein modification is directly related to the cancer. Thus, this technique was novel in providing the first comprehensive look at N-glycan modification and associated protein characterization in a site-directed manner for HCC, and in the future, other types of cancer as well.

# **Chapter 3: Changes in Glycosylation of Hepatocellular Carcinoma via MALDI-IMS**

As noted previously, the first aim of the project was to examine the N-linked glycan changes in HCC as compared to previously studied changes in serum. The following chapter details the process and results of examining hepatocellular carcinoma tissues for the cancerous regions specifically, looking at the changes in N-glycosylation, as well as identifying trends in a patient sample cohort. This bulk of this chapter was included as part of a publication published in the Journal of Proteome Research in October of 2018. Contributions in writing, experiments, data analysis, and final approval were done by myself, with intellectual and minor editorial contributions from others on the publication [161].

### **3.1. Abstract**

Hepatocellular carcinoma (HCC) remains as the fifth most common cancer in the world and accounts for more than 700,000 deaths annually. Changes in serum glycosylation have long been associated with this cancer but the source of that material is unknown and direct glycan analysis of HCC tissues has been limited. Our laboratory previously developed a method of *in situ* tissue based N-linked glycan imaging that bypasses the need for microdissection and solubilization of tissue prior to analysis. We used this methodology in the analysis of 138 HCC tissue samples and compared the N-linked glycans in cancer tissue with either adjacent untransformed or tissue from patients with liver cirrhosis but no cancer. Ten glycans were found significantly elevated in HCC tissues as compared to cirrhotic or adjacent tissue. These glycans fell into two major

classes, those with increased levels of fucosylation and those with increased levels of branching with or without any fucose modifications. In addition, increased levels of fucosylated glycoforms were associated with a reduction in survival time. This work supports the hypothesis that the increased levels of fucosylated N-linked glycans in HCC serum are produced directly from the cancer tissue.

### **3.2. Introduction**

Changes in N-linked glycosylation are known to occur with the development of many cancers, including hepatocellular carcinoma (HCC) [119, 122, 129, 242-251]. In previous work, serum was examined for protein glycoforms that are altered in liver cancer and significant alterations in serum N-linked glycosylation with the development of HCC were documented [151, 180, 181, 191, 252-257]. Specifically, the alterations are increased levels of alpha-1,3 and alpha-1,6 linked fucosylation found on bi, tri and tetra-antennary glycans and to a lesser extent alterations in high mannose and tetra-antennary glycans [151, 180, 181, 191, 252-257]. Importantly, many of these changes are now being developed as serum-based biomarkers of HCC. However, the origins of these glycans in human HCC are unknown and glycan analysis of tissue is complicated by the mixing of different cell types and the loss of protein during processing. To address these limitations, the lab has previously developed a method of tissue-based glycan imaging that allows for both qualitative and quantitative in situ N-linked glycan analysis on tissue using matrix-assisted laser desorption/ionization

mass spectrometry imaging (MALDI-MSI) [195, 258-260]. This method bypasses the need for microdissection and solubilization of tissue prior to analysis. When matrix is applied across the tissue section, desorption can be targeted to specific “points” in a pattern and the data rasterized. The resulting spectra can then be used to generate two-dimensional heat maps of hundreds of glycans directly from the surface of a tissue section. These molecular maps display the relative abundance and spatial distribution of these molecules. Thus, MALDI tissue profiling has the power to link the molecular detail of mass spectrometry with molecular histology, generating mass spectra correlated to locations within a thin tissue section. This method is becoming a robust technique for the analysis of glycan *in situ* [195, 258-266]. In this study, we used this methodology in the analysis of two tissue microarrays (TMA). The first TMA consisted of 48 HCC tissue samples, 22 cirrhotic tissue samples and 5 healthy control tissue samples. The second TMA consisted of 90 HCC tissue samples and 90 control adjacent tissue samples. MALDI glycan imaging has identified 10 glycans that were significantly increased in the HCC TMA samples when compared to cirrhotic tissue (TMA #1) or to non-transformed adjacent tissue (TMA #2). These glycans fell into two major classes, those with increased levels of fucosylation and those with increased levels of branching without any fucose modifications. The relevance of this finding to serum based biomarkers and the potential prognostic role of these glycans is discussed.

### **3.3. Materials and Methods**

#### **3.3.1. Materials**

Trifluoroacetic acid, Harris-modified hematoxylin, and  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) were obtained from Sigma Aldrich (St. Louis, MO). HPLC grade methanol, ethanol, acetonitrile, xylene, hydrogen peroxide and water were obtained from Fisher Scientific (Pittsburgh, PA). Tissue Tack microscope slides were purchased from Polysciences Inc (Warrington, PA). Citraconic anhydride and SafeClear II was purchased from Thermo Scientific (Bellefonte, PA). Recombinant Peptide N-Glycosidase F (PNGase F) from *Flavobacterium menigosepticum* was obtained, expressed, and purified as previously described [267], but is also available commercially as PNGase F Prime™ from Bulldog Bio (Portsmouth, NH). Universal Antigen Retrieval Reagent was purchased from R&D Systems (Minneapolis, MN).

#### **3.3.2. Tissues and Tissue Microarrays**

Normal and hepatocellular carcinoma whole liver tissue samples were purchased from ProSci Inc. (Poway, CA) and cirrhotic whole liver tissue was purchased from BioChain (Newark, CA). All tissue microarray (TMA) slides were purchased from US Biomax (Rockville, MD) as unstained formalin fixed paraffin embedded (FFPE).

The first TMA (Catalog Number: BC03117) contained 80 cores. Forty-eight cases of HCC with a history of Hepatitis B virus (HBV) infection, five cases of cholangiocellular carcinoma with a history of Hepatitis B virus (HBV) infection,



22 cases of liver cirrhosis with a history of Hepatitis B virus (HBV) infection and five normal hepatic tissue cores. These cores were 1.5 mm in diameter and 5  $\mu$ m thick. For the purpose of this study, the cholangiocellular carcinoma tissue was included in any analysis.

The second TMA slide (Catalog Number: HLiv-HCC180Sur-04) contained 90 cases of HCC with tumors ranging from stage 1 (early) to 4 (late) and grades G1 (well-differentiated) to G3 (poorly differentiated). All HCC tissues had matched un-transformed adjacent tissue. Along with this, survival data and pathology diagnosis was included for each case. The cores were cut at a 1.5 mm diameter and a thickness of 4  $\mu$ m.

### **3.3.3. Washes for Deparaffinization**

As described previously [259], FFPE TMA slides were heated at 60°C for 1 hr and cooled to room temperature prior to deparaffinization. The slides were washed with xylene to remove the paraffin and then rehydrated using a series of water and ethanol washes. Antigen retrieval was performed using citraconic anhydride (Thermo Scientific) as the buffer and placed in a vegetable steamer for 30 minutes. The buffer was then cooled to room temperature and buffer exchange was performed to replace the slides in 100% water. Finally, the slides were desiccated until dry.

### 3.3.4. Enzymatic Digestion and Matrix Deposition

A M3 TM-Sprayer™ Tissue MALDI Sample Preparation System (HTX Technologies, LLC) was used to spray 0.5 mL of 0.1 µg/µl aqueous solution PNGase F as previously described [259]. Following the spray, the slides were placed in a humidified chamber and incubated at 37°C for 2 hours. Slides were then desiccated and dried prior to matrix application. The matrix used was  $\alpha$ -cyano-4-hydroxycinnamic acid (0.042 g CHCA in 6 mL 50% acetonitrile/49.9% water/0.1% TFA) and sprayed using the same M3 TM-Sprayer.

### 3.3.5. N-Glycan Imaging using MALDI-IMS

The slides were analyzed for released N-glycan ions using a Solarix dual source 7T FTICR mass spectrometer (Bruker Daltonics,  $m/z$  500-5000) with a SmartBeam II laser operating at 1000 Hz and with a laser spot size of 25 µm. 200 laser shots were collected for each pixel, with a time domain of 512K. This resulted in a resolving power of 160,000 at  $m/z$  400. A total of 23,145 positions were collected for TMA #1 and 44,533 positions collected for TMA #2. Afterwards, the data was analyzed using FlexImaging 4.0 (Bruker Daltonics) and SCiLS Lab (Bruker Daltonics, version 2017b) to create images and determine regions of differentially expressed glycans, all normalized to total ion current. A signal to noise (S/N) ratio of 9 was used and peaks were manually picked within FlexImaging 4.0. The resulting glycans were given composition using an in-house database based on collected  $m/z$  values and checked against the database from GlycoWorkbench based on  $m/z$  and composition [268]. Possible

and likely structures for visual representation were built using GlycoWorkbench as well.

### **3.3.6. Lectin Histochemistry**

The tissue slides were deparaffinized by using PROTOCOL SafeClear II clearing agent, then rehydrated in a series of ethanol washes at 3 minutes per each step (100%, 90%, 70%) and fully hydrated in deionized water for 5 minutes. Endogenous peroxidase activity was blocked using 3% hydrogen peroxide, followed by a 92°C heated antigen retrieval using Universal Antigen Retrieval Reagent (Dako, Carpinteria, CA). The slides were then fixed with 4% formaldehyde solution followed by a permeabilization step using 0.5% IGEPAL CA-630. Following the permeabilization step, for blocking non-specific binding, the slides were blocked once again with serum-free protein block (Dako), supplemented with Streptavidin/Biotin blocking solution to block endogenous biotin, biotin receptors, and streptavidin binding sites in tissues (Vector Laboratory, Burlingame, CA). Streptavidin horseradish peroxidase ready to use solution (Vector Laboratory) was used to detect biotinylated recombinant *Aleuria aurantia* N224Q (rAAL N224Q) lectin bound to the tissue, and visualization was further developed using 3,3'-diaminobenzidine (DAB) Chromogen (Dako). The N224Q lectin is a modified *Aleuria aurantia* lectin with increased binding to core fucosylated glycan (Herrera et al., manuscript submitted). Lectin was applied for 1 hour at room temperature at a concentration of 0.5 µg/mL in background reducing antibody diluent (Dako, Carpinteria, CA). After incubation, slides were

washed with TBS (pH 7.6) for 5 minutes in room temperature and repeated three times. Data on this lectin can be found in [193, 269, 270]. Finally, the slides were counterstained with Harris-modified hematoxylin (Fisher Scientific, Hampton, NH) for increased visualization. Annotation was done digitally using Aperio ImageScope (Leica Biosystems, Buffalo Grove, IL) for positive pixel signal algorithm based on lectin staining.

### **3.3.7. Statistical analysis**

For all peaks (m/z values) mean intensity values were determined for each individual TMA spot. To facilitate statistical analysis, original data was transformed by log based on 10. Further descriptive statistics and statistical inference are all based on the log-transformed data.

To compare difference of glycan between HCC tissues and cirrhotic tissue, we applied t-test or Wilcoxon rank sum test, appropriately on data distribution. For tumor tissue and its adjacent tissue comparison, paired t-test or Wilcoxon rank sum test was also selected based on glycan data distribution. Two-sided hypothesis test was selected, p-value less than 0.05 was considered statistically significant. Receiver Operator Curves (ROC) curves were constructed, area under curve (AUC) was considered as discriminant ability, standard error of AUC was derived from bootstrap.

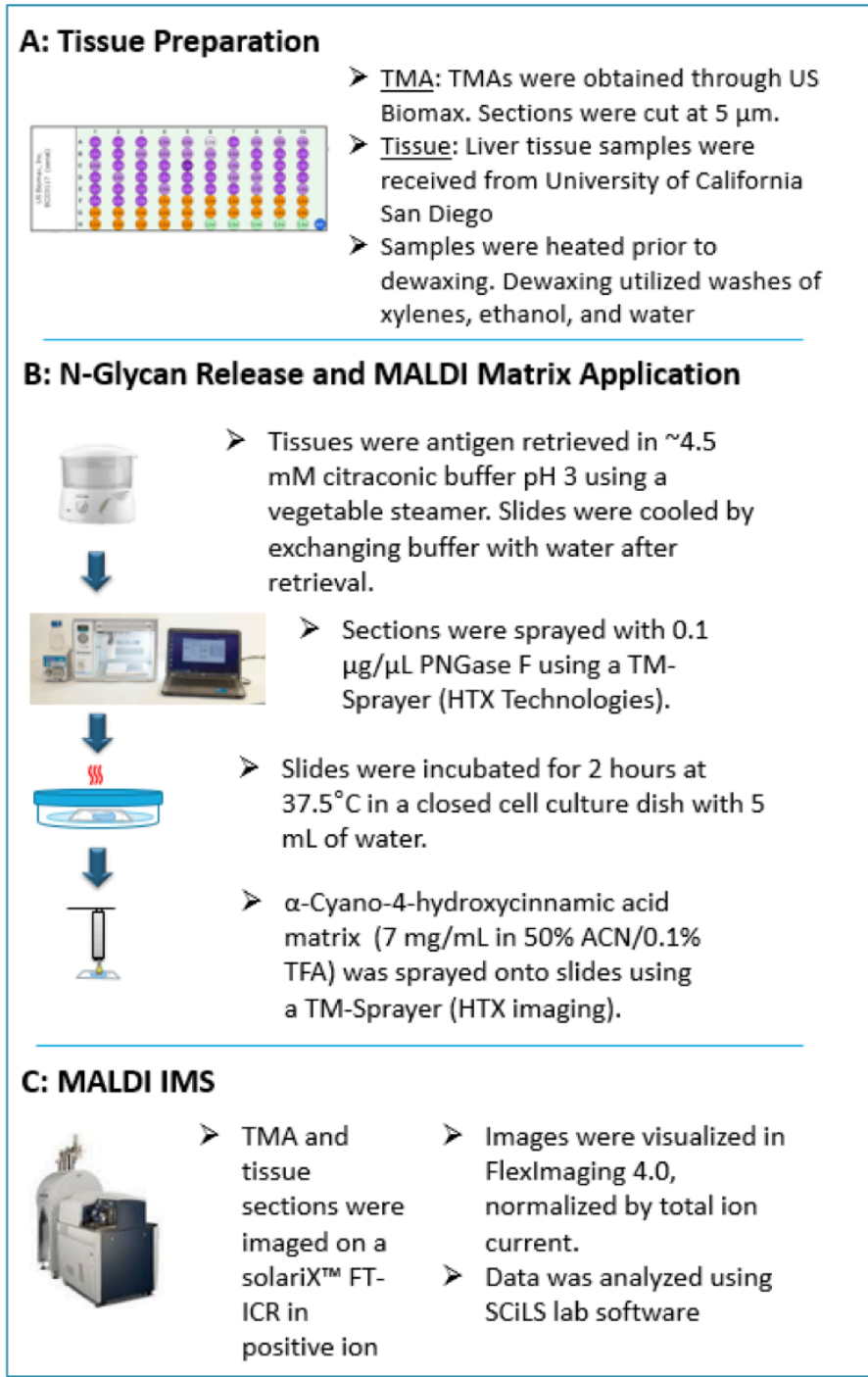
For survival analysis, the median of the specific glycan was used as a cutoff line to classify patients who were above the median as being in the high group and the rest as the low group (choosing mean as cutoff derived similar

results, because the mean and median were similar in 3 concerned glycans). We plotted the Kaplan-Meier survival curves of the high and low group, and log-rank test was applied to check survival difference between two groups. The Cox Proportional hazard model was used for further analysis.

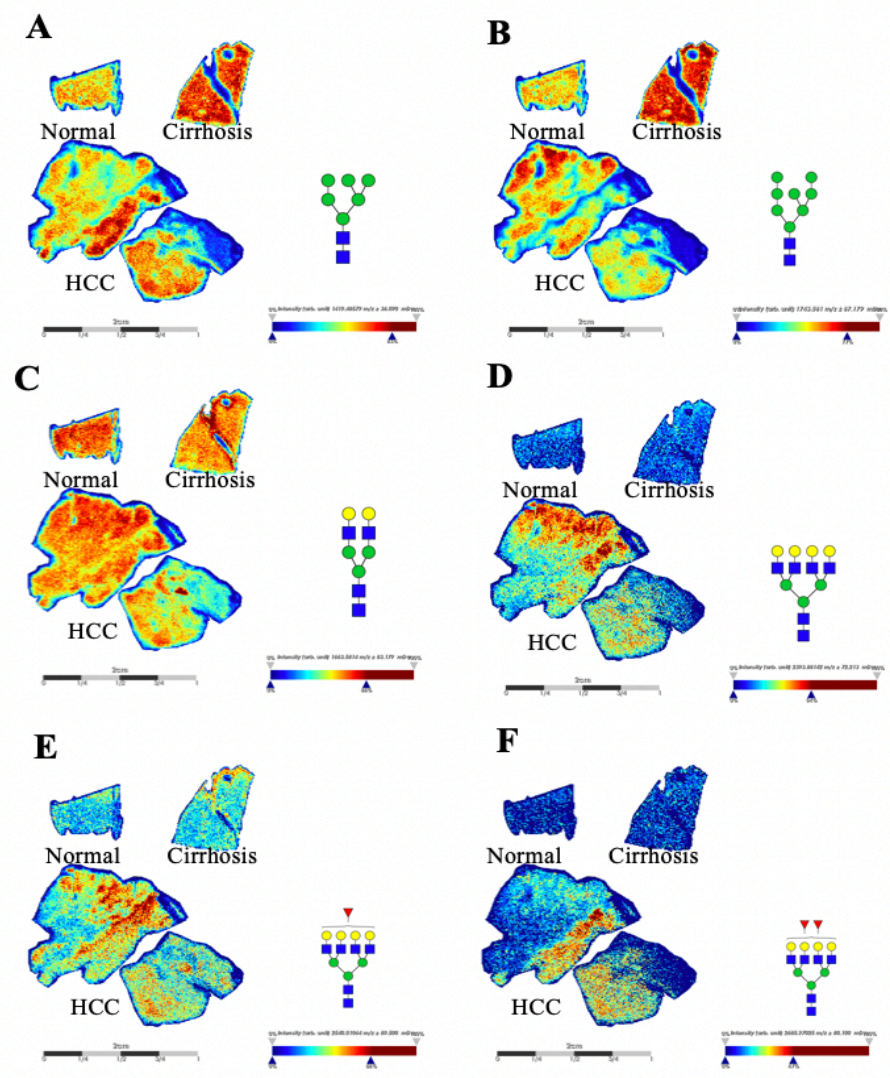
### **3.4. Results**

#### **3.4.1. Increased Complex and Fucosylated N-Glycans in Liver Tissues**

We, and others, have previously correlated changes in glycosylation in the serum of individuals with the development of HCC [151, 180, 181, 191, 252-257] [122, 176, 184, 187, 188, 190, 271-274]. To determine the glycan changes that occur directly in HCC tissue we utilized a MALDI based glycan imaging methodology [259] to examine the N-linked glycans that alter with the development of HCC. In our initial experiments, we examined five sections of HCC tissue obtained from patients with hepatitis B virus (HBV) - associated HCC, three sections of normal tissue and three sections of cirrhotic tissue. Figure 11 shows the workflow of the tissue analysis and Figure 12 shows the results of a representative normal, cirrhotic and HCC tissue. In this figure, specific N-linked glycans are shown and their relative abundance presented via a heat map of individual glycan intensities across each tissue, where blue is low abundance and red is high abundance. Figure 13 shows the same sections stained with hematoxylin and eosin staining in both a 1X and 10X magnification confirming diagnosis. In Figure 12, three N-linked glycans that were found in all tissues and

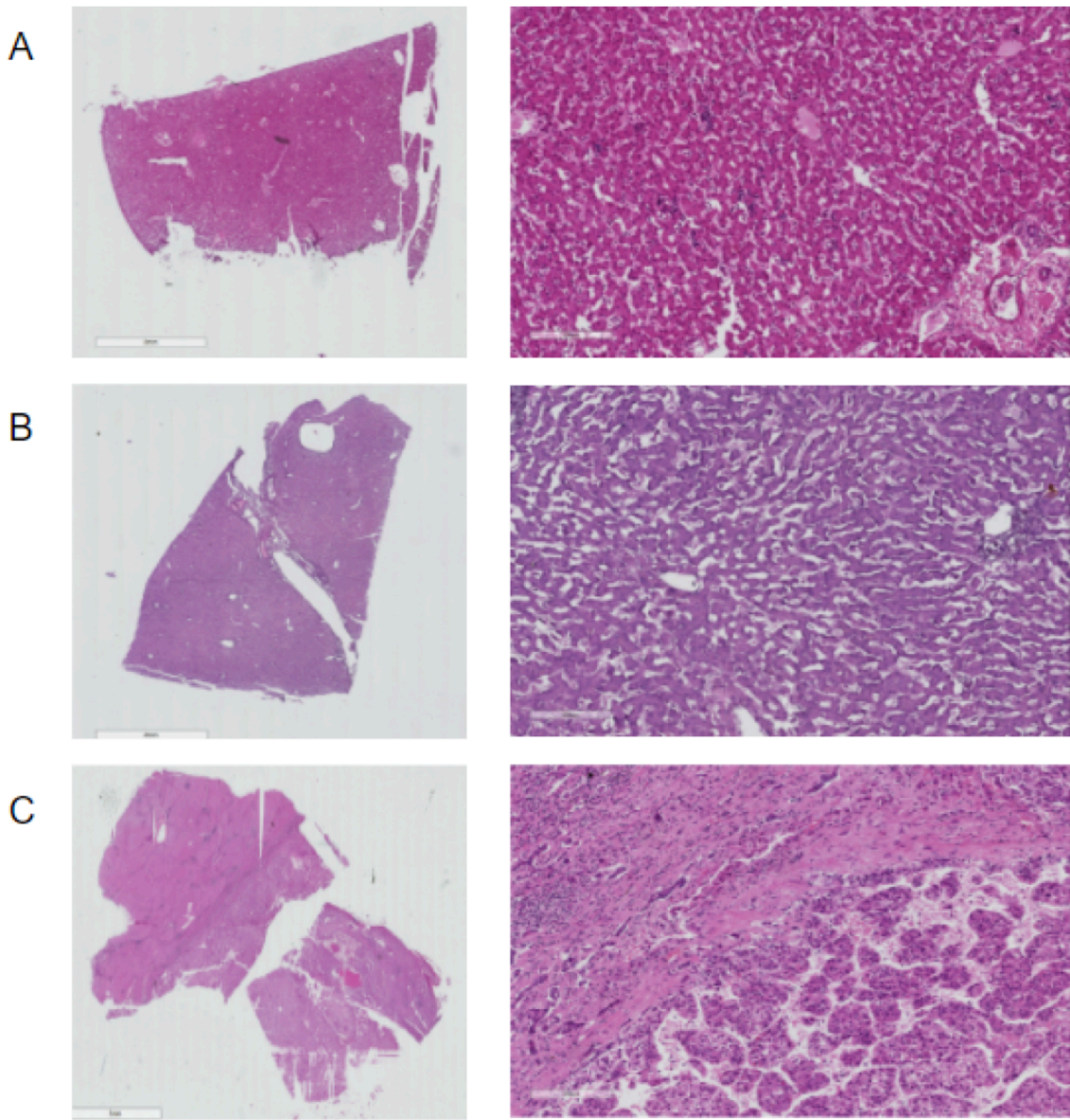


**Figure 11: Workflow of Tissue-based Glycan Analysis.** Generalized workflow for slide prep and MALDI IMS imaging



**Figure 12: Detection of N-Glycans in Normal, Cirrhotic and HCC Tissues.**

While certain glycans are found in all tissues (A-C), some glycans are found predominantly in the HCC tissue (C-E). Images were acquired with 150  $\mu\text{m}$  raster step size on a Bruker 7T solariX XR ICR FTMS system. Ion intensities are normalized to the TIC of each ion across the tissue. Color scale bars are included and autocorrected for the range of intensities plotted. For glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose<sub>68</sub>



**Figure 13: Hematoxylin and Eosin Staining of Varying Tissue Types.**

H&E staining at 1x and 10x magnification for normal (A), cirrhotic (B), and HCC tissue (C).



three N-linked glycans that were elevated in the HCC as compared to the normal or cirrhotic tissue are presented. As Figure 12A-C shows, high mannose glycans such as Man 7, Man 8 (Figure 12A&B) and simple bi-antennary type glycans without substantial fucosylation (Figure 12C) can be found equally in normal, cirrhotic and HCC tissue. In contrast, glycans such as a tetra-antennary glycan without fucosylation (Figure 12D) or tetra-antennary glycan with single (Figure 12E) and multiple fucose residues (Figure 12F) are found predominantly in the HCC tissue as compared to the normal or cirrhotic tissue. It is noted that these mass values (glycan) were not observed without the application of PNGase F (data not shown).

#### **3.4.2. Analysis in Human Liver Tissue Microarray Set**

To determine whether these N-glycan changes seen in the HCC tissue could be observed in a larger set of tissue samples, we examined two independent tissue microarrays (TMAs), one consisting of 48 HCC, 22 cirrhotic, and 5 normal tissue cases and another TMA consisting of 90 samples with HCC and 90 tissue samples of the adjacent untransformed tissue. Clinical and patient information for these samples are provided in Table 2 and 3. Figure 14 shows both TMAs, demonstrating the imaging data received from the workflow. Table 4 presents a list of all the potential glycans found in both TMAs. Elevations in specific glycans was determined by examining the mean intensity values of each glycan structure in the HCC tissue and in the cirrhotic tissue for

**Table 2: Patient Characteristics for TMA #1.** Characteristics of TMA #1 including number, diagnosis, etiology, age, gender, grade and stage

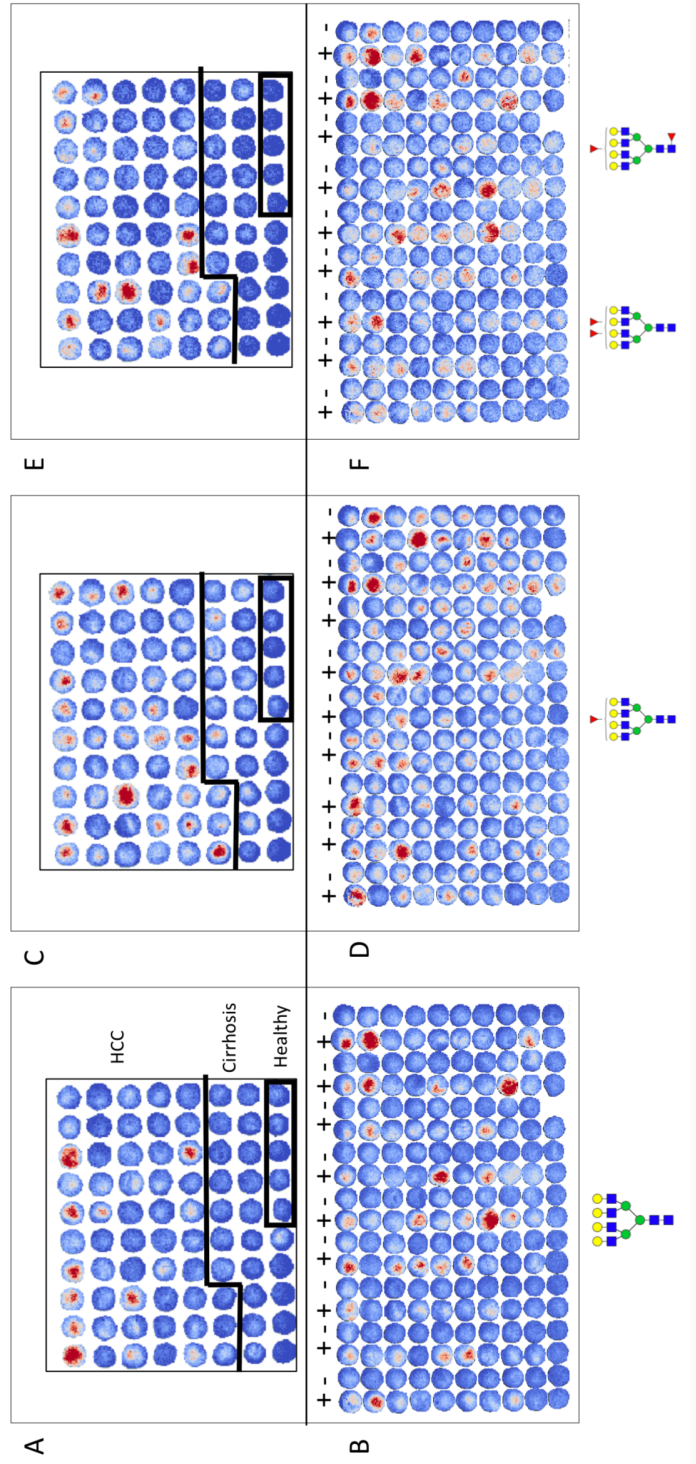
<b>Disease Diagnosis</b>	<b>HCC<sup>1</sup></b>	<b>Cirrhosis<sup>2</sup></b>
Number	48	22
Etiology% (HBV/HCV/other) <sup>3</sup>	100/0/10	100/0/0
Age (mean, SD) <sup>4</sup>	51.6 (9.9)	51.0 (8.4)
Gender (M/F) <sup>5</sup>	40/8	19/3
Grade <sup>6</sup>	2.09 (0.66)	---
Stage <sup>7</sup>	2.67 (0.55)	---

**1&2)** Disease diagnosis was determined by MRI (1) or by liver biopsy (2). **3)** For Etiology: HBV, hepatitis B virus; HCV, hepatitis C virus; other, liver disease consisting of cryptogenic liver disease or alcohol induced liver disease. **4)** Mean age of groups. **5)** Mean gender of the groups. **6)** Grade of tumor as defined by American Joint Committee on Cancer (AJCC) TNM surgical grading. Mean value and standard deviation indicated. **7)** Tumor staging information as defined by AJCC TNM clinical staging. Mean value and standard deviation indicated.

**Table 3: Patient Characteristics for TMA #2.** Characteristics of TMA #2 including number, diagnosis, etiology, age, gender, grade, stage, and survival time

<b>Disease Diagnosis</b>	<b>HCC<sup>1</sup></b>	<b>Normal Adjacent<sup>2</sup></b>
Number	90	90
Etiology% (HBV/HCV/other) <sup>3</sup>	Unknown	Unknown
Age (mean, SD) <sup>4</sup>	53.9 (10.0)	53.9 (10.0)
Gender (M/F) <sup>5</sup>	81/9	81/9
Grade (mean, SD) <sup>6</sup>	2.21 (0.47)	---
Stage (mean, SD) <sup>7</sup>	2.40 (0.75)	---
Survival time <sup>8</sup> (with range)	33.3 (1-80)	---

**1&2)** Disease diagnosis was determined by MRI (1) or by liver biopsy (2). **3)** For Etiology: HBV, hepatitis B virus; HCV, hepatitis C virus; other, liver disease consisting of cryptogenic liver disease or alcohol induced liver disease. **4)** Mean age of groups. **5)** Mean gender of the groups **6)** Grade of tumor as defined by AJCC TNM surgical grading. Mean value and standard deviation indicated. **7)** Tumor staging information as defined by AJCC TNM clinical staging. Mean value and standard deviation indicated. **8)** Mean survival time in months with range given.



**Figure 14: Representative Imaging Data from both TMA Datasets.** Representative image data collected for both TMA sets (first TMA with independent samples is on top with the second TMA with HCC and adjacent tissue is on bottom) showing three different glycan structures: A) 2393.840 m/z, B) 2539.957 m/z and C) 2685.969 m/z. For the first TMA, the HCC tissue is indicated at top with the cirrhotic samples in the middle and the healthy tissue samples in the box at the bottom right corner. For the second TMA, the + above each column represents the HCC tissue and the consecutive – column represents the matched normal adjacent tissue section. The proposed glycan is presented at the bottom of each panel

**Table 4: Master List of N-Linked Glycans.** Table showing all found N-glycans, including mass, error, and structure

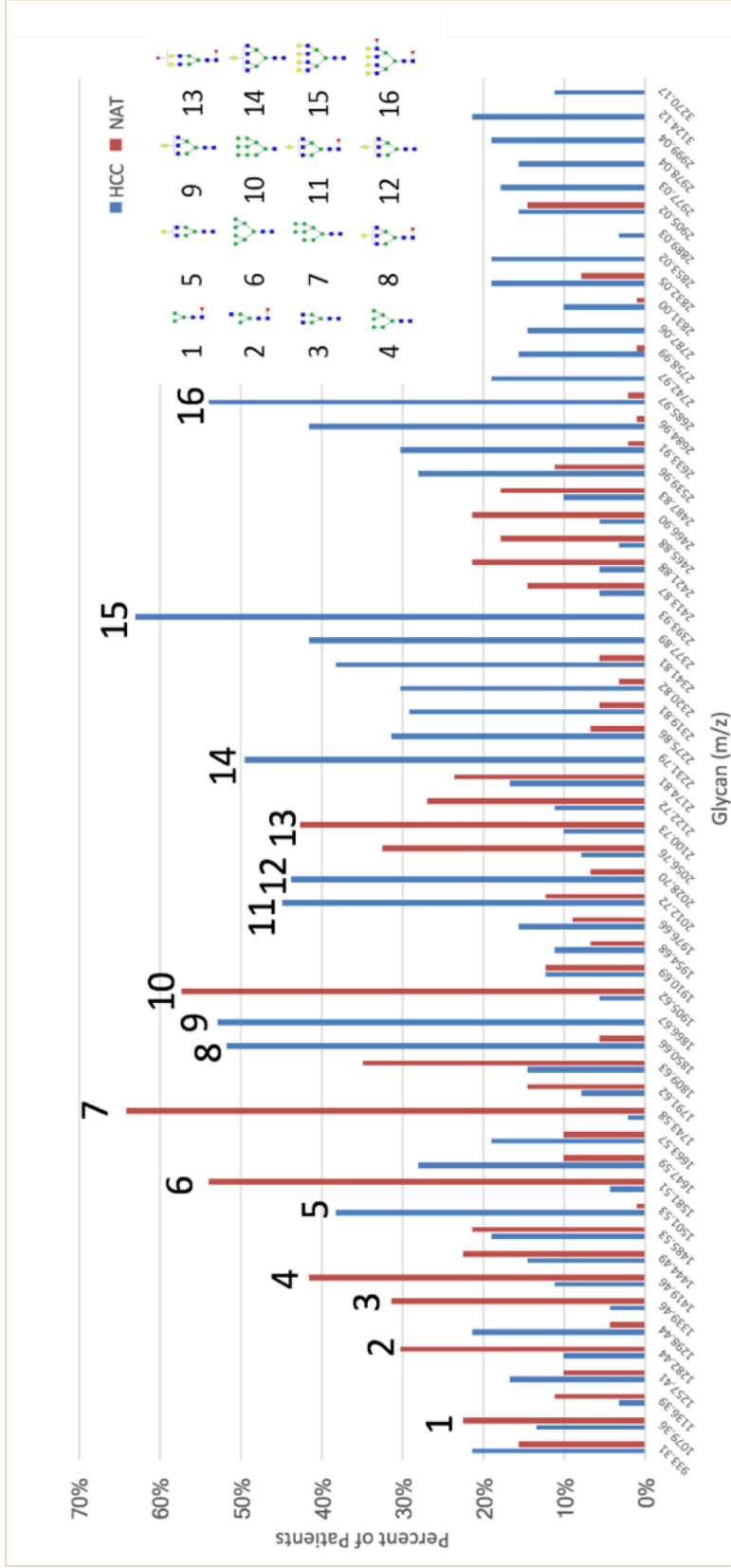
Centroid [m/z]	± [Da]
933.31458	0.04667
1079.3618	0.05397
1087.3769	0.05437
1095.3658	0.05477
1118.374	0.03355
1136.3933	0.03409
1239.4149	0.03718
1257.4145	0.03772
1282.438	0.03847
1298.4354	0.03895
1339.4629	0.04018
1419.4633	0.04258
1444.4933	0.04333
1460.5039	0.04382
1485.5274	0.04457
1501.5314	0.04505
1562.5343	0.04688
1581.5054	0.04745
1645.5608	0.04937
1647.5892	0.04943
1663.567	0.04991
1708.5922	0.05126
1743.5755	0.05231
1791.6204	0.05375

Centroid [m/z]	± [Da]
1809.6332	0.05429
1850.6558	0.05552
1866.6663	0.056
1905.6195	0.05717
1910.6938	0.05732
1954.6818	0.05864
1976.6594	0.0593
2012.7242	0.06038
2020.654	0.06062
2028.7019	0.06086
2056.7582	0.0617
2100.7265	0.06302
2122.7173	0.06368
2158.8017	0.06476
2174.8057	0.06524
2231.7913	0.06695
2275.8647	0.06828
2289.7287	0.06869
2319.8133	0.06959
2320.8242	0.06962
2341.8106	0.07025
2377.8885	0.07134
2393.84	0.07182
2413.868	0.07242

Centroid [m/z]	± [Da]
2421.8766	0.07266
2465.8777	0.07398
2466.8952	0.07401
2487.8291	0.07463
2539.957	0.0762
2633.9132	0.07902
2684.9645	0.08055
2685.9689	0.08058
2742.9741	0.08229
2758.9912	0.08277
2787.0606	0.08361
2830.9955	0.14155
2832.0464	0.08496
2853.0197	0.08559
2889.0319	0.08667
2905.0211	0.08715
2977.031	0.08931
2978.0434	0.08934
2999.0364	0.08997
3035.1603	0.09105
3051.0658	0.09153
3124.121	0.09372
3197.1893	0.09592
3270.1723	0.09811

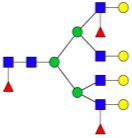
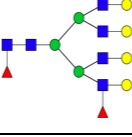
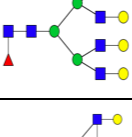
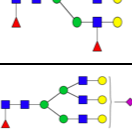
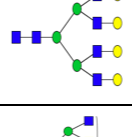
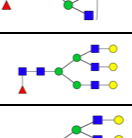
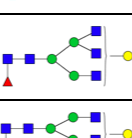
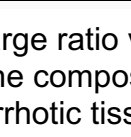



TMA #1 or the un-transformed adjacent tissue in TMA #2. A mean intensity value increase of 1.5-fold in the HCC sample as compared to the appropriate control tissue was considered elevated. Similarly, if the intensity was 0.5 times that of the appropriate control tissue, the structure levels were decreased. Figure 15 highlights the observation that many of the 61 observed glycans were seen in 20-40% of the HCC tissue samples and often in less than 5% of the control tissue.

Two glycans that were observed as elevated in over 50% of the TMA samples, were a tetra-antennary glycan (glycan #15 in Figure 15) and a tetra-antennary glycan with two fucose residues (glycan #16 in Figure 15). As this family of glycan - tetra-antennary glycan with and without fucosylation - were observed in many of the samples, we further examined the level of this family of glycans in the TMAs. The level of the tetra-antennary glycan lacking fucose (Figure 14A), the tetra-antennary glycan with a single (Figure 14B) and double fucose residues (Figure 14C) in both the TMAs are shown. As before, darker red colors represent a higher intensity for the specific glycan while more blue tones represent less intensity. The mean values of signal intensities for specific glycans found in the HCC tissue as compared to the cirrhotic tissue (in TMA#1) or between the HCC tissue and the adjacent non-transformed tissue (TMA#2) were compared. Table 5 presents glycans (selected by lowest  $p$  value) that were significantly elevated ( $p < 0.05$ ) in the HCC tissue as compared to the cirrhotic (TMA #1) or adjacent non-transformed tissue (TMA #2) as well as a glycan that was not altered in the HCC tissue. A master list of all N-glycan  $m/z$  values is



**Figure 15: Total Patient Glycan Upregulation.** Percentage of patients with upregulated glycans for HCC and control tissues

**Table 5: Glycans Altered in HCC Versus Cirrhosis or Adjacent Tissue.**

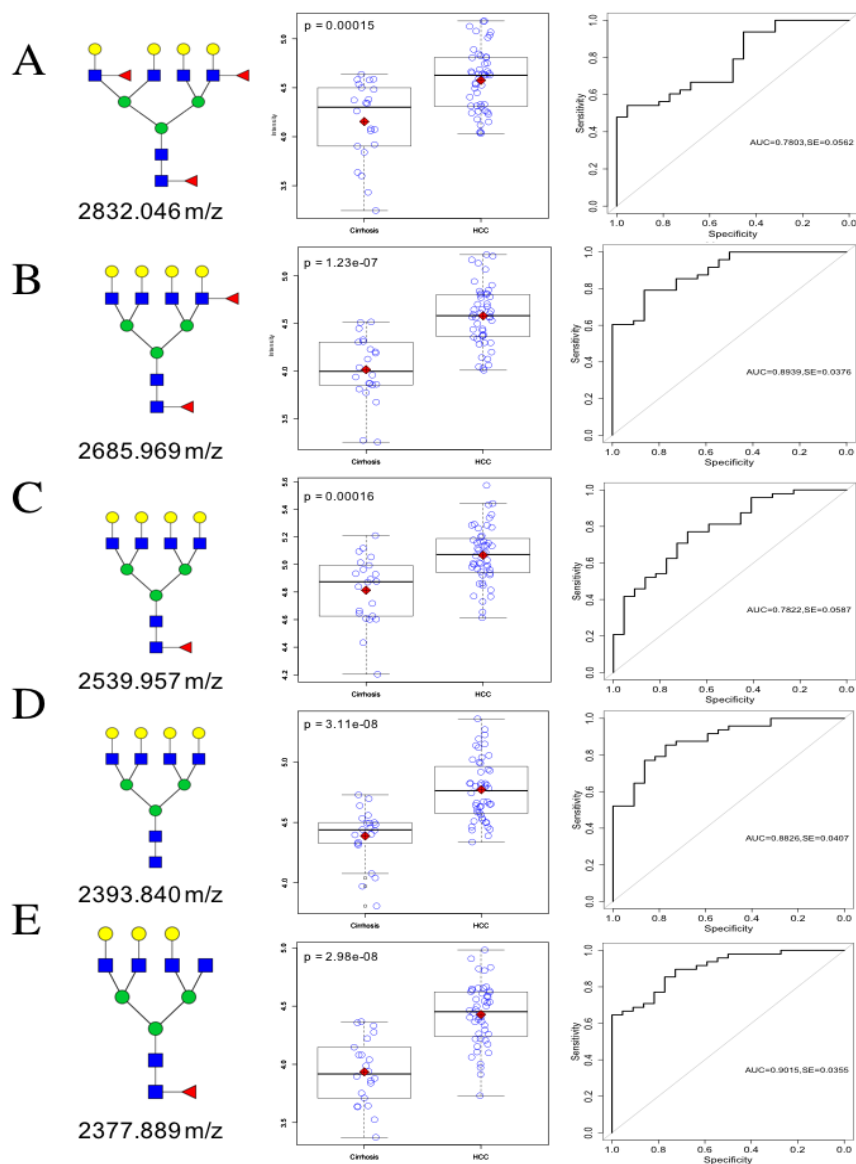
Observed m/z <sup>1</sup>	Proposed Glycan <sup>2</sup>	P value TMA 1 <sup>3</sup>	P value TMA 2 <sup>4</sup>
2832.046		0.00015	0.2105
2685.969		1.23e-07	7.59e-15
2539.957		0.00016	0.02739
2466.895		0.0195	0.002727
2465.878		0.2427	0.000654
2393.840		3.11e-08	9.33e-16
2377.889		2.98e-08	6.71e-11
2174.806		0.03208	0.3902
2012.724		0.000281	0.000105
1850.656		0.000103	1.71e-09
1647.589		0.00633	0.00571

1) Observed mass to charge ratio value 2) The proposed glycan structure based upon the m/z value. 3) The composition of the identified M/Z value. 4) P value comparing the HCC to cirrhotic tissue. Analysis by students T-test. 5) P value comparing the HCC to adjacent tissue.

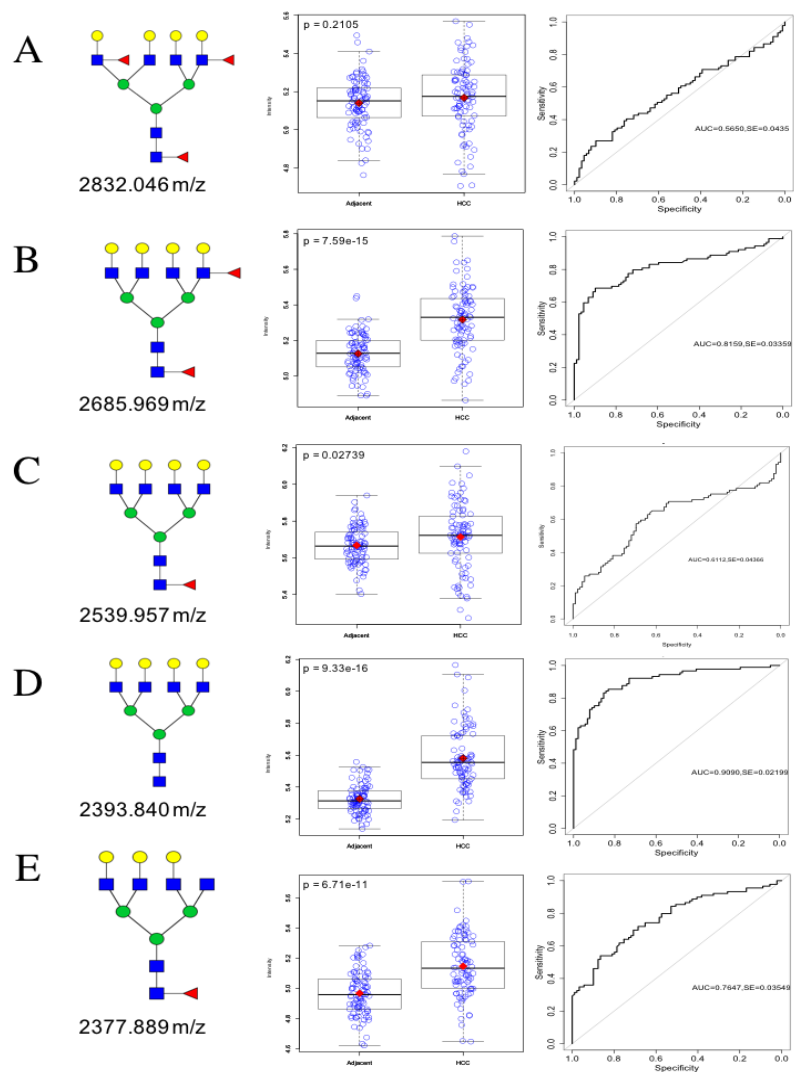


provided (Table 4). As Table 5 shows, nine out of ten glycans that were elevated in the HCC tissue were fucosylated glycan, The glycans with the lowest  $p$  value were tetra-antennary glycans, with or without fucosylation. To further explore the branched and fucosylated glycome in these two TMAs, we examined the five tetra-antennary glycan that were altered in the TMA's by scatter plot and by AUROC analysis. Figure 16 shows these data for TMA #1 and Figure 17 shows these data for TMA #2. As these figures show, alterations in specific tetra-antennary glycoforms could be observed in both TMAs. For example, a glycan at  $m/z$  2685.969, proposed as a di-fucosylated tetra-antennary glycan, was elevated in HCC tissue in both TMA #1 and TMA #2 (Figure 16B and 17B). Similarly, the tetra-antennary glycan ( $m/z$  2393.840) devoid of fucosylation was also altered in both TMAs (Figure 16D and 17D). In contrast, TMA #1 had greater alterations in a tetra-antennary glycan with three fucose residues, as compared to TMA #2 (Figure 16A and 17A). Other versions of the tetra-antennary glycan family were also observed in both TMAs (Figure 16C and 17C).

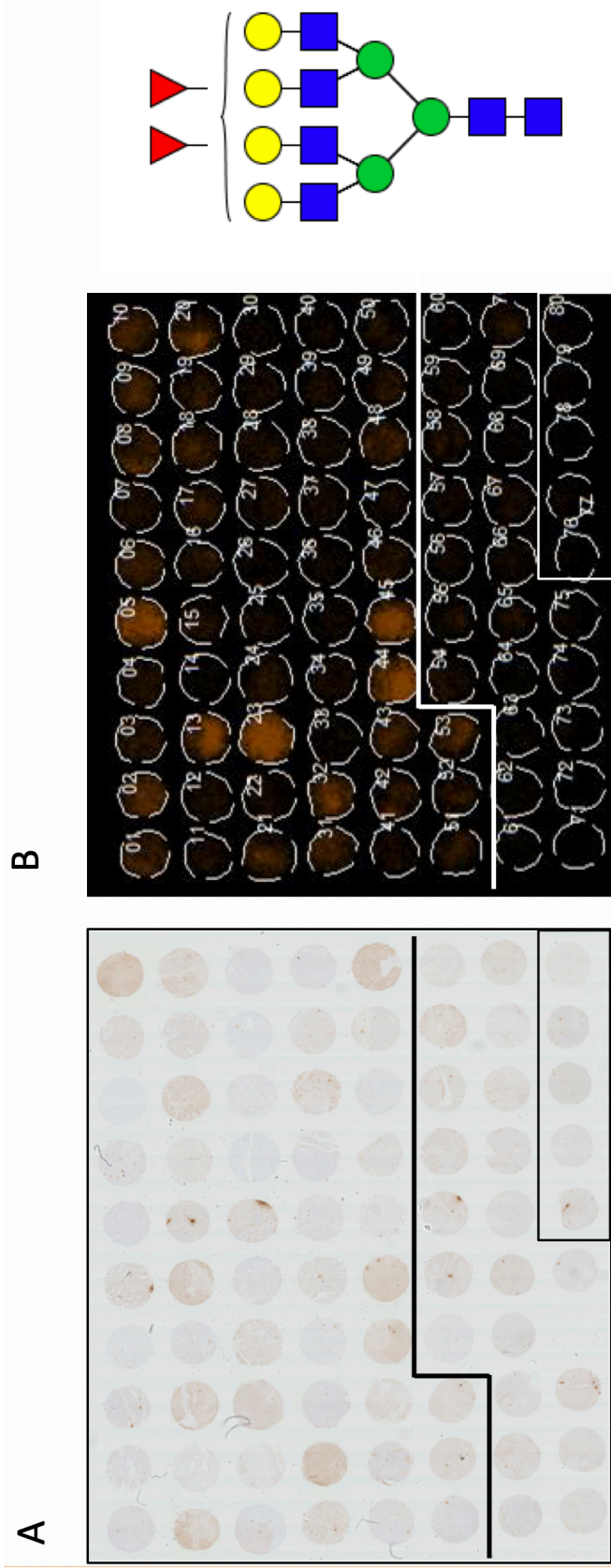
Increased fucosylation seen by MALDI-MSI was further confirmed by lectin histochemistry. Figure 18 shows the lectin histochemistry staining for one of the TMAs using a recombinant *Aleuria aurantia* lectin (AALN224Q) lectin which has enhanced binding to branched and core alpha 1,6 lined fucosylated glycan and reduced binding to alpha 1,2 linked fucose [275, 276]. Figure 18, shows a side by side comparison of the lectin histochemistry and the MALDI imaging for one of the most prominent fucosylated glycan ( $m/z$  2685.969; see Table 5), supporting the classification as fucosylated structures.



**Figure 16: Analysis of Human Liver TMA #1.** Proposed glycan structure, log transformed intensity scatter plot with the red diamond indicating mean and associated p-value, and Receiver Operating Characteristic (ROC) curve with AUC value for select structures. In panels A-E, analysis was done comparing HCC versus cirrhotic samples. A, B, C, and E utilized a student t-test for their p-value while D utilized a Wilcoxon Rank Sum Test.



**Figure 17: Analysis of Human Liver TMA #2.** Proposed glycan structure, log transformed intensity scatter plot with the red diamond indicating mean and associated p-value, and Receiver Operating Characteristic (ROC) curve with AUC value for select structures. In panels A-E, analysis was done comparing HCC versus cirrhotic samples. A, B, C, and E utilized a student t-test for their p-value while D utilized a Wilcoxon Rank Sum Test.

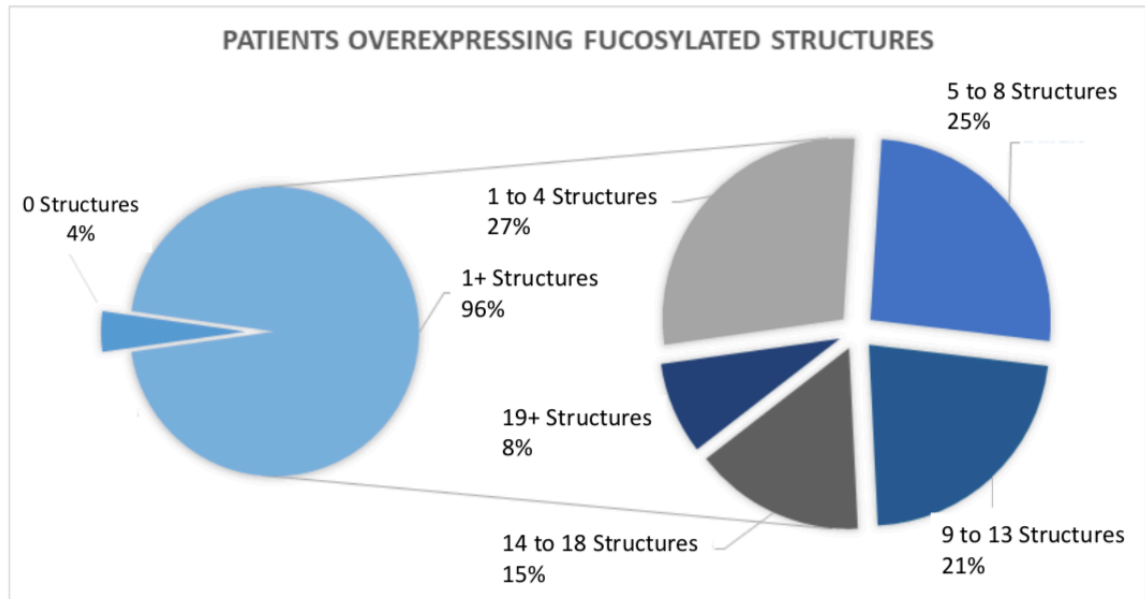


**Figure 18: Lectin Stain Compared to MALDI-IMS Data.** Lectin staining for fucosylated structures and corresponding MALDI IMS data for a fucosylated glycan structure

As fucosylation was a prominent feature of the altered glycan in HCC, with 33/61 of the proposed glycan structures containing fucose, fucosylation levels were further explored. To accurately determine the elevated levels of the fucosylated glycans seen between HCC and adjacent tissue, we examined the difference between the adjacent and HCC tissue of each individual patient in the matched tissue set TMA and determined the percentage of patients with elevated levels of each of these glycans in both their HCC tissue and their matched normal adjacent tissue. Elevation was again determined by using mean intensity values of these glycans in both the HCC and non-transformed adjacent tissue and if the value was 1.5 times that of normal levels, the patient was considered to have elevated levels of that glycan. As Figure 19 shows, 96% of patients had increased levels of at least one fucosylated structure. Those patients were then categorized into the number of these highly branched and/or fucosylated structures they were presenting, with patients demonstrating increased levels in anywhere from one fucosylated structure to all 33 found within the TMA.

### **3.4.3. Association of Specific N-Glycans with Survival**

For the patient-matched TMA (TMA #2), survival data were available allowing for the determination of an association between glycan and outcome. This was done for the three major glycans observed in the HCC tissue: a tetra-antennary glycan (m/z 2393.840), a tetra-antennary glycan with a single fucose (m/z 2539.957) and a tetra-antennary glycan with two fucose residues (m/z 2685.969). Patients with glycan expression greater than the median level in all



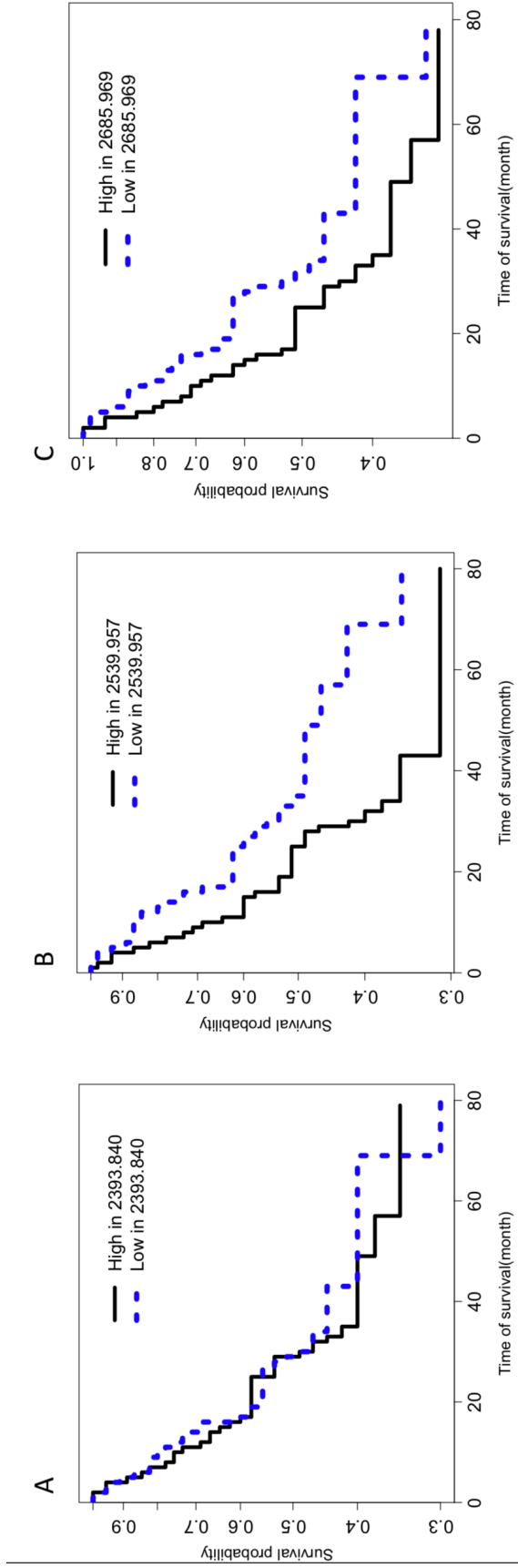
**Figure 19: Patients Demonstrating Elevated Levels of Fucosylated Glycans.**

A total of 33 fucosylated glycans were found elevated in the patient-matched TMA. Comparing HCC to the un-transformed adjacent tissue, a 1.5x relative intensity increase in HCC tissue was used to classify patients as elevated. Of the 89 patients able to be analyzed, 96% (85 patients) demonstrated elevated levels of at least one of these fucosylated structures (left). Of these 85 patients, they were further classified into varying classes based on the number of fucosylated structures they had elevated levels for. 27% (23 patients) had elevated levels of one to four fucosylated structures, 25% (21 patients) had elevated levels off five to eight structures, with 8% of patients showing elevated levels of 19 or more of these fucosylated structures with one patient showing elevated levels all 33 fucosylated structures found.

tissue (both tumor and adjacent normal) were classified as high. There was no association with these glycans between the levels observed in the normal adjacent tissue and patient outcome (data not shown). In addition, as Figure 20A shows, the level of the m/z 2393.840 glycan was also not associated with patient outcome. The mean time of survival was 29 months in those with high or low levels of the m/z 2393.840 glycan. In contrast, as Figures 20B and 20C show, patients with high levels of the m/z 2539.957 or m/z 2685.969 glycans were associated with shorter survival times. For the m/z 2539.957 glycan, those with high levels had a median survival time of 25 months, while patients with low levels of this glycan had a median survival time of 35 months. Similarly, for the m/z 2685.969 glycan, those with high levels had a median survival time of 25 months, while patients with low levels of this glycan had a median survival time of 32 months. Cox proportional hazard model analysis showed patients with one unit increase of the m/z 2685.969 glycan would enhance the hazard(risk) about 3-fold,  $p=0.0334$ . One unit increase of the m/z 2539.957 glycan would increase hazard about 8-fold,  $p=0.0078$ . There was no association between these glycans and stage or grade of HCC (data not shown).

### **3.5. Discussion**

Alterations in glycosylation have been long observed with HCC [185, 277-281]. Much of this work was shown in serum, with little analysis directly in the HCC tissue itself or has been analyzed following dissection of tissue and the mixing of the multiple hepatic (and non-) cell types. Here we utilized MALDI-IMS



**Figure 20: Survival Plots for Branched and Fucosylated Glycans.** Kaplan–Meier Survival plots for glycan at m/z 2393.840 (A) 2539.957 (B) and 2685.968 (C) from the 90 patient TMA. Survival time is in months.



glycan imaging to identify the glycans that occur directly in 138 HCC patient tissue samples. In our analysis of both TMA's, there were 61 glycans that were found to be upregulated in at least one HCC tissue sample (See Table 4). In addition, there were 10 glycans that were significantly ( $p < 0.05$ ) increased in at least 30% of the HCC tissue samples as compared to either cirrhotic or adjacent tissue (see Figure 15).

Our previous MALDI glycan imaging developmental work had utilized a small number of HCC tissue samples and a 16 patient HCC TMA [195]. In that study, alterations in both branching and fucosylation were observed but the sample size was too small to determine the significance of the changes detected. Here, we have extended that work through an analysis of a larger number of samples and also with the association with outcome data regarding the observed glycans.

Surprisingly, only two major classes of glycan were observed in HCC tissue as compared to either cirrhotic tissue or adjacent untransformed tissue. The first was a tetra-antennary glycan structures and the second was an increase in the level of fucosylation. It is noted that the tetra-antennary glycan was only observed in HCC tissue and not in adjacent tissue or cirrhotic tissue. The tetra-antennary glycan is formed through the action of alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A (MGAT5), which has been associated with many cancers through mutations of the telomerase reverse transcriptase (hTERT) [282] and through activation of the Ras/Raf pathway [66].

The second major alteration observed in the HCC tissue was increased fucosylation. This glycan change has been observed in the serum for many years, but a clear understanding of where this material derives was not known. However, glycan analysis of tumor derived material was unable to identify fucosylation as being increased in HCC [198, 201]. This was most likely the result of the method used, which involved homogenization of tissue and mixing of cell types. In contrast, by using the MALDI glycan imaging method we were able to observe increased levels of fucosylation on independent sample sets. Most often on tetra-antennary glycan but also to a lesser extent on bi-antennary and tri-antennary glycan. Indeed, there is now significant evidence that transformed hepatocytes are the source cells for serum fucosylated proteins. Recent work showed that as hepatocytes undergo an epithelial–mesenchymal transition (EMT), they increase the genes, such as alpha-1,6-fucosyltransferase gene (*FUT8*), which are involved in fucosylation [199]. This is consistent with lung cancer, where *FUT8* increased as a direct result of an EMT [200]. In addition, a recent report has indicated that HCC downregulates miR-122 and leads to the upregulation of *FUT8* [15]. It is also noted that the deletion of *FUT8* in a mouse model inhibits chemical induced HCC by the down regulation of cancer associated signaling pathways [156, 186]. Together, this data suggests very strongly that fucosylation originates from the cancer itself and prior analytical glycan tools were not able to detect this change within the tumor. In addition, over 95% of the HCC samples analyzed had increased levels of one or more fucosylated glycan, while normal adjacent tissue did not, supporting the

hypothesis that fucosylation is an event associated with the general act of cellular transformation.

In addition, we observed only a few sialylated structures by MALDI-glycan imaging and it is possible that these 1) were not detected by our method or 2) we had sialic acid loss following ionization. It is also highly likely that both things are true and methods to stabilize sialic acids will be required for analysis of these structures [262].

While the identity of the proteins containing these changes are unknown, several proteins have been characterized as containing the glycans shown to be up-regulated in HCC tissue. For example, we have recently shown that low molecular weight (LWM) kininogen contains fucosylated tetra-antennary glycan and that this protein can act as a serum biomarker of HCC [155]. Additionally, serum fucosylated haptoglobin and fucosylated fibronectin have been shown to contain branched fucosylated tetra-antennary glycan in HCC [272].

Lastly, heterogeneity was observed in the glycans associated with HCC and it is assumed that this most likely is the result of the underlying genetic heterogeneity found with the disease [15]. In conclusion, I have shown that two major glycan changes are associated with HCC, increased branching and increased fucosylation. Hopefully, in the future, these glycan changes can be exploited for the early detection of HCC and potentially in the treatment of HCC.

While many of these changes are novel and exciting developments in terms of N-glycan modifications for HCC, there are still challenges associated with these changes. For example, while we were able to identify an increase in

fucosylation across the board for HCC tissue specifically, one challenge of imaging mass spectrometry is that the technique is only able to identify predictive structural components based on  $m/z$ , not linkage. Where this is extremely relevant is in looking at core fucosylation specifically. Imaging mass spectrometry techniques are unable to accurately identify whether the fucosylation modification is core or outer arm, and as discussed previously, FUT8 and core fucosylation are fundamentally more related to cancer progression and HCC in the literature. The lectin staining does support the hypothesis that fucosylation is occurring in the HCC tissue (Figure 18). Therefore, this led us to further investigate the possibility of determining core versus outer arm fucosylation while still maintaining all the benefits afforded to us through imaging.

# **Chapter 4: Determination of Core Fucosylation using MALDI-IMS**

Alluded to in the previous chapter, the inability to determine core versus outer arm fucosylation in an imaging mass spectrometry experiment caused a disconnect in associating previously studied N-glycan changes in hepatocellular carcinoma with new-found data. Therefore, it was imperative to develop a new method for the determination of core versus outer arm fucosylation while maintaining all the benefits of an imaging mass spectrometry experiment. The chapter below was partially included as part of a publication that was accepted with revisions in the Journal of Proteome Research in March of 2020. Contributions in writing, imaging experiments, data analysis, and final approval were done by myself, with intellectual, minor editorial, and enzyme production and validation were done by others on the publication (Hongyan Liang, Richard Drake, and Anand Mehta).

#### **4.1. Abstract**

Specific alterations in N-linked glycans, such as core fucosylation, are associated with many cancers and other disease states. Because of the many possible anomeric linkages associated with fucosylated N-glycans, determination of specific anomeric linkages and site of fucosylation (i.e., core versus outer arm) can be difficult to elucidate. A new MALDI mass spectrometry imaging workflow in formalin-fixed clinical tissues is described using recombinant Endoglycosidase F3 (Endo F3), an enzyme with a specific preference for cleaving core fucosylated N-glycans attached to glycoproteins. In contrast to the broader substrate enzyme Peptide-N-Glycosidase F (PNGaseF), Endo F3 cleaves between the two

core N-acetylglucosamine residues at the protein attachment site. On tissue, this results in a mass shift of 349.137 a.m.u. for core fucosylated N-glycans when compared to N-glycans released with standard PNGaseF. Endo F3 can be used singly and in combination with PNGaseF digestion of the same tissue sections. Initial results in liver and prostate tissues indicate core fucosylated glycans associated to specific tissue regions while still demonstrating a diverse mix of core and outer arm fucosylated glycans throughout all regions of tissue. By determining these specific linkages while preserving localization, more targeted diagnostic biomarkers for disease state is possible without the need for microdissection or solubilization of the tissue.

#### **4.2. Introduction**

It is well established that many aspects of the molecular development and progression of cancer are directly linked to changes in glycosylation [119-121, 123-129]. In most cases, glycan analysis has been done with serum and not directly from the cancer tissue itself [122, 134-149]. Serum is often used as it is easily obtained, but it is limited in that it is comprised of dilute levels of tumor-derived material. Thus, direct tissue analysis is preferred. However, the mixing of different cell types, and the loss of protein during processing complicate glycan analysis of tissue, often leading to misleading data and misrepresentation of tumor specific analysis. To combat this, the lab has previously developed a method of tissue-based glycan imaging that allows for both qualitative and quantitative *in situ* N-linked glycan analysis on tissue using matrix-assisted laser

desorption/ionization mass spectrometry imaging (MALDI IMS) [258]. This method was co-developed in 2013 by the Drake and Mehta laboratories, and has continued to evolve [195, 259, 260] to allow for better analysis of sialylated glycan [262] and for the simultaneous analysis of glycan and protein [261]. However, a major limitation of the MALDI-TOF imaging methods is the inability to obtain true structural and linkage information of a PNGase F released glycan. To address this limitation, we began to examine other enzymes that may allow for more structural information via imaging mass spectrometry.

In 1982, a novel glycosidase preparation from *Flavobacterium meningosepticum*, designated Endo- $\beta$ -N-acetylglucosaminidase F was described [283] and found to include three distinct endoglycosidase activities, termed Endo F1, Endo F2, and Endo F3[284]. These three endoglycosidases cleave the  $\beta$ (1-4) link between the two core GlcNAc of asparagine-linked glycans, but have specificities for distinct oligosaccharide structures [285]. For example, Endo F1 cleaves high mannose) and hybrid structures, but not complex oligosaccharides and core fucosylation of hybrid structures reduces the rate of cleavage by 50-fold. Endo F2 cleaves primarily complex glycan with core fucosylation having little impact upon glycan cleavage. In contrast, Endo F3 has no activity on oligomannose and hybrid molecules it has a reported 400-fold increase in activity toward core fucosylated structures as compared to tri-antennary structures at a pH of 4.5, thus reducing the amount of non-specific N-glycan cleavage [286]. With this in mind, Endo F3 was applied to different MALDI IMS workflows alone or in conjunction with PNGase F. This workflow would allow for the structural



characterization of core fucosylated glycans in tissue while maintaining the localization of N-glycans in tissue.

Fucosyltransferase 8 (FUT8), the only known enzyme responsible for core fucosylation, has been implicated in a variety of settings including non-small cell lung cancer, melanoma, and hepatocellular carcinoma demonstrating an increase in invasion and metastasis for patients with elevated levels of FUT8 or core fucosylated N-glycans[99-101]. These previous studies show the importance of determining core fucosylation as opposed outer arm fucosylation in terms of N-linked glycosylation and the clinical relevance of this methodology described below.

### **4.3. Materials and Methods**

#### **4.3.1. Cloning, Expression, and Purification of Endo F3**

The cDNA fragment encoding the Endo F3 gene was amplified by PCR from the genomic DNA of *Elizabethkingia meningoseptica* (UniProtKB – P36913) without the N-terminus signal sequence. Additionally, a His tag (x10) was added to its C-terminus. Amplified DNA fragments were cloned into pQE-60 by NcoI/BlnI (Genscript, Piscataway, NJ). The constructed plasmid, pQE-60-Endo F3-10xHis, was transformed into BL21 (DE3). The transformants were cultured in LB broth supplemented with 100 µg/ml Ampicillin. Cultures were grown at 37°C until the cells reached an  $A_{600nm}$  of about 0.5, 0.5 mM IPTG were added to the culture to induce protein overproduction at 20°C. The next day, the cells were harvested by centrifugation. The cell pellets were re-suspended in PBS with

added Pierce protease inhibitor tablets (ThermoFisher Scientific, Waltham, MA), stored at -20°C. Omnicleave endonuclease (Lucigen Corporation, Middleton, WI) and MgCl<sub>2</sub> were added to thawed cell suspension. The cell suspension was incubated at room temperature for at least one hour with rocking. The cells were lysed using a French Press (GlenMills Inc., Clifton, NJ) per the manufacturer's instructions. The cell lysis was applied to HisTrap FF (GE Healthcare, Pittsburgh, PA) and washed with 20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole (pH 7.4). Bound His-tagged protein was eluted with a gradient from 150 to 500 mM imidazole in 20 mM sodium phosphate, 0.5 M NaCl (pH 7.4). The purified Endo F3 was desalted and concentrated with 20 mM Tris-HCl, 50 mM NaCl (pH 7.5) using Spin-X UF Concentrator (10kDa; Corning). The protein purity was confirmed using SDS-PAGE.

#### **4.3.2. In-Solution Digestion by Endo F3**

Human Fetuin-A (Assaypro, St. Charles, MO) or RNase B (New England BioLabs) were incubated with Endo F3 at an enzyme-to-protein-ratio of 1:5 (w/w) at 37°C for 3 hours. For our purposes, 1µg of Endo F3 was added to 5µg of protein at a pH of 4.5.

#### **4.3.3. Glycan Sequencing**

Human Fetuin-A was run on SDS-PAGE gel, stained and cut out. The gel pieces were alkylated in the dark for 30 minutes with iodoacetamide, fixed in a solution of 10% methanol 7% acetic acid for one hour, washed in acetonitrile,

followed by subsequent steps of 20 mM ammonium bicarbonate (pH 7.0) and acetonitrile before being dried in a speed-vac. PNGase F (PNGase F Prime™, N-Zyme Scientifics, Doylestown, PA) or Endo F3 was diluted with corresponding buffer and allowed to absorb into and cover the gel pieces, then incubated overnight at 37°C. The glycans were eluted from the gel pieces by sonication in Milli-Q water, dried down and labeled with a 2AB dye as previously described [287]. The labeled glycans were subsequently enriched from free 2AB dye using paper chromatography and filtered using PTFE syringe filter unit. Fluorescently labeled glycans were then separated on normal phase Waters Alliance HPLC system as previously described [287]. Samples were further digested with Sialidase for calculation of glucose unit (GU) value and compared to GlycoStore database [288].

#### **4.3.4. On-Slide Tissue Preparation and Imaging**

Multiple formalin-fixed paraffin-embedded (FFPE) blocks of tissue were obtained for optimization and analysis. Tissue microarray (TMA) slides were purchased from US Biomax (Rockville, MD) while all other tissue blocks (prostate, cervix, and liver) were provided by the Medical University of South Carolina Biorepository and Tissue Analysis Shared Resource (Charleston, SC). The FFPE blocks were sectioned on to slides at 5µm then prepped for imaging as previously described [289]. Briefly, the slides were washed and deparaffinized by heating at 60°C for one hour, then washed sequentially in xylene, a dilution of ethanol, and water. The slides then underwent antigen retrieval using citraconic

anhydride and placed in a steam chamber for 30 minutes. Finally, buffer exchange was performed and desiccated. Enzyme was then applied to the slides using an M3 TM-Sprayer™ Tissue MALDI Sample Preparation System (HTX Technologies, LLC) at 0.1 µg/µl. PNGase F was sprayed in HPLC water while Endo F3 was sprayed in a solution of 87 µM acetic acid (pH 4.43) for better efficiency. The slides were then placed in a humidity chamber and incubated at 37°C for 2 hours then desiccated. Finally, matrix was applied ( $\alpha$ -cyano-4-hydroxycinnamic acid, 0.042g CHCA in 6 mL 50% acetonitrile/49.9% water/0.1% trifluoroacetic acid) using the same M3 TM-Sprayer.

As previously described, tissues were analyzed via imaging N-glycans using both a MALDI FTICR mass spectrometer (Solarix Dual Source, 7T, Bruker Daltonics,  $m/z$  500-5000). The data was then analyzed and visualized using FlexImaging 5.0 and SCiLS Lab 2017b (Bruker Daltonics). Finally, glycans were built and validated against the database in GlycoWorkbench, as well as built for graphical interpretation [258, 268].

#### **4.3.5. N-Glycan Removal**

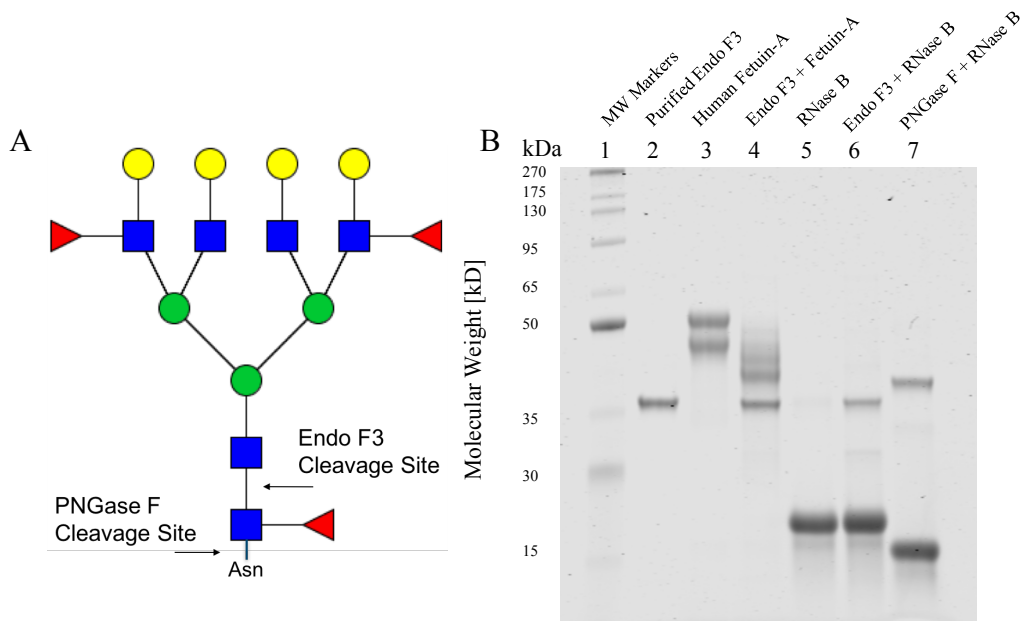
In cases where F3 was applied first, glycans were collected from the slide and analyzed as previously described [290]. Briefly, the slides were placed in 100% ethanol for removal of matrix, then placed in series of dilutions of ethanol (95% and 70%). Next, the slides were placed in a high pH cleaning solution (10mM Tris, pH 8.98), HPLC grade water, then a low pH cleaning solution (citric acid buffer, pH 3), then HPLC grade water again. The slides were then

desiccated and dried. Following the cleaning, the tissues were then prepped for PNGase F application following the same tissue preparation and glycan imaging protocol as previously described (40), however the dewaxing and antigen retrieval steps were omitted, beginning with enzyme application on the tissue.

#### **4.4. Results**

##### **4.4.1. In-solution analysis of Endo F3 activity on N-Linked Glycans**

The deglycosylation activity of the purified recombinant Endo F3 was tested initially using two well characterized glycoproteins, RNase B and Fetuin-A, to confirm the activity of Endo F3 acting on core fucosylated glycans only (Figure 21). Human Fetuin-A is a circulating plasma glycoprotein with two N-linked and three O-linked carbohydrate side chains [291]. The heterogeneity of Fetuin-A is mainly due to extensive modification with variable amounts of sialic acids; some less abundant glycoforms were found to be core-fucosylated [292]. RNase B is a well characterized glycoprotein from bovine pancreas that only contains non core-fucosylated high mannose N-glycans attached to a single N-linked glycosylation site [293]. As shown by the SDS-PAGE, the recombinant Endo F3 will cleave Fetuin-A but not RNase B as shown by the band shift on the gel, which is consistent with Endo F3 reported sensitivity and specificity. In contrast, treatment with PNGase F leads to a band shift of RNase B. This supports the claim that we can differentially cleave glycans on proteins based on the composition of the glycans attached to them, specifically ignoring high mannose glycans that don't contain a core fucose modification.



**Figure 21: SDS-PAGE Analysis of N-Glycans by PNGase F or Endo F3. A)**

Cartoon description of Endo F3 vs. PNGase F cleavage on core fucosylated

N-linked glycans. For glycans, red triangle, fucose; blue square, N-

acetylglucosamine; green circles, mannose; yellow circles, galactose. B) SDS-

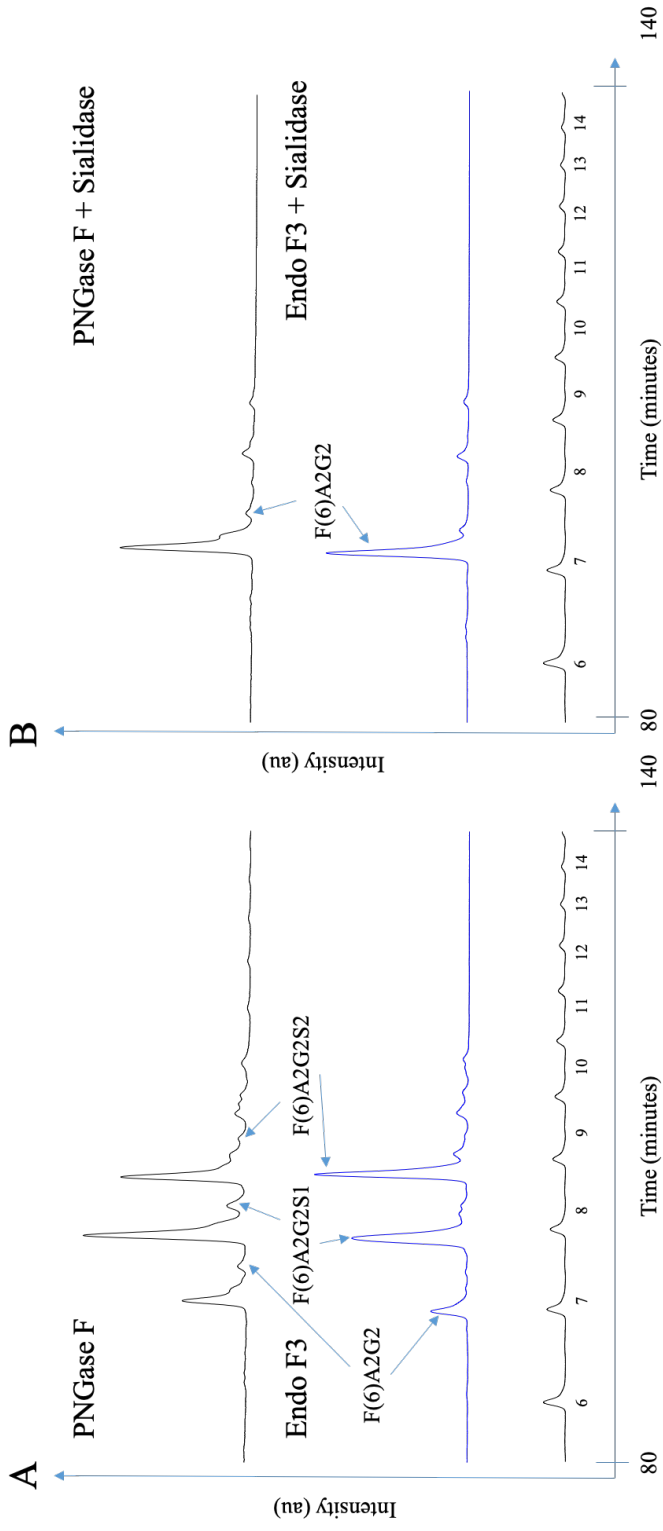
PAGE analysis of Endo F3 and PNGase F digestion of human Fetuin-A or

RNase B.

The glycan profile of Fetuin-A was also investigated by normal phase HPLC. The chromatograms are shown in Figure 22. A standard curve using the homopolymer dextran was used to convert the elution time into glucose units is shown at the bottom of the figure. Among PNGase F released glycans, sialic acid removal simplified the profiles and further treatment with bovine kidney fucosidase (result not show) removed peaks at GU 7.70, representing a biantennary glycan with a core  $\alpha$ -1,6-linked fucose (F(6)A2G2) that only contributed 2.2% of the total glycan profile. On the other hand, the three major glycans released by Endo F3 are all core fucosylated biantennary with variable amounts of sialic acids that represented 83.0% of the total glycan profile; with the removal of sialic acids, the three species were combined into one peak at GU 7.20 which is F(6)A2G2\*, considering that with Endo F3 digestion, one GlcNAc and the core  $\alpha$ -1,6-linked fucose was left on the protein as opposed to the cleavage at the asparagine residue for PNGase F (Figure 21A).

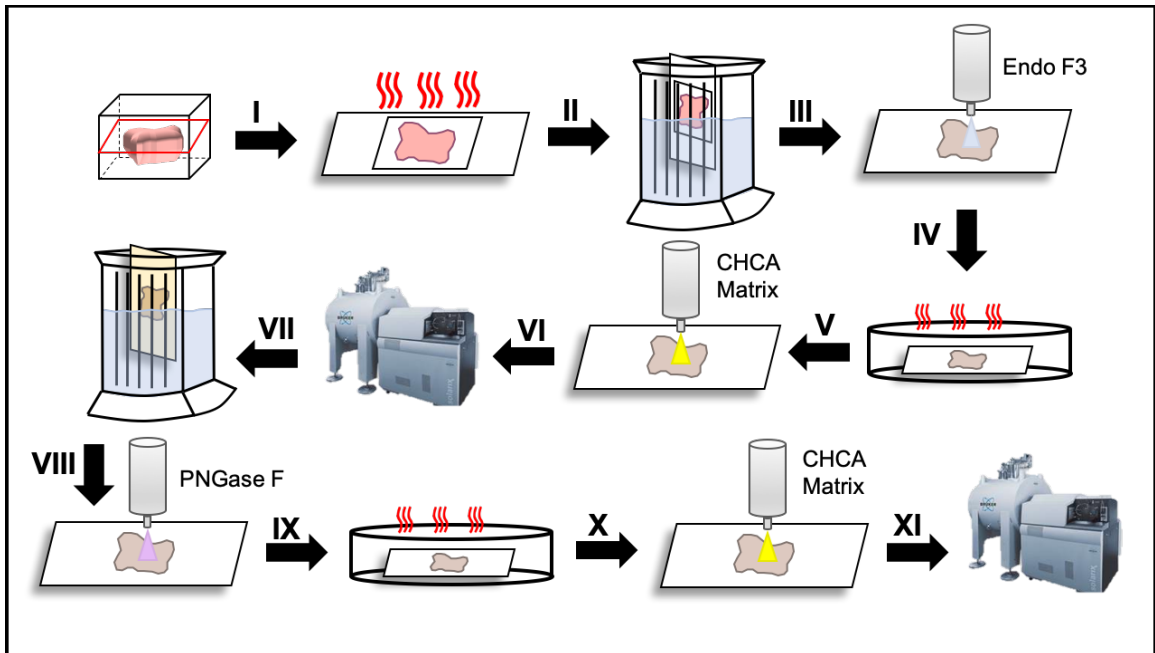
#### **4.4.2. On-Tissue Analysis of Endo F3 Digestion using MALDI Imaging Mass Spectrometry**

Keeping the conserved GlcNAc and fucose residue in mind, we then applied the enzyme to the well-established tissue imaging protocol as described above (Figure 23). With the differential cleavage of Endo F3 as compared to PNGase F, we see a mass shift of 349.137 m/z for core fucosylated glycans. When applied, we see the downward shift in the mass spectra of core fucosylated N-glycans while effectively prohibiting cleavage of N-glycans that do



**Figure 22: HPLC Analysis of Fetuin A N-Glycans Following Digestions.** A) N-linked glycans sequencing of human Fetuin-A released by PNGase F or Endo F3, B) followed by sequential sialidase digestion.

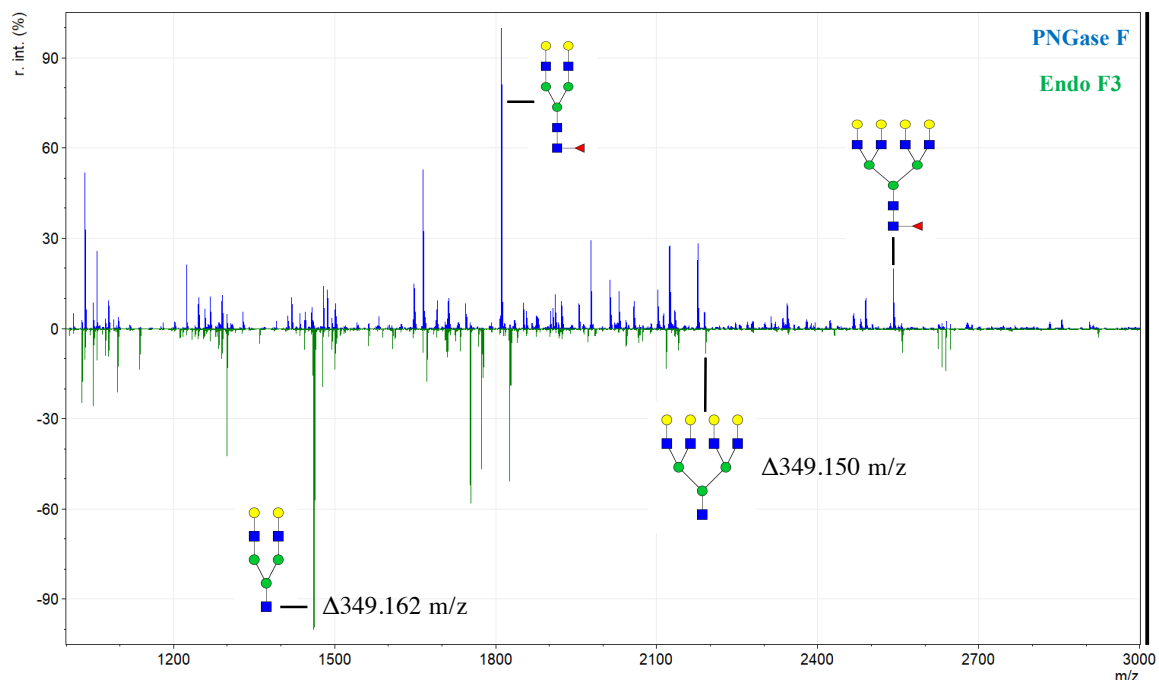




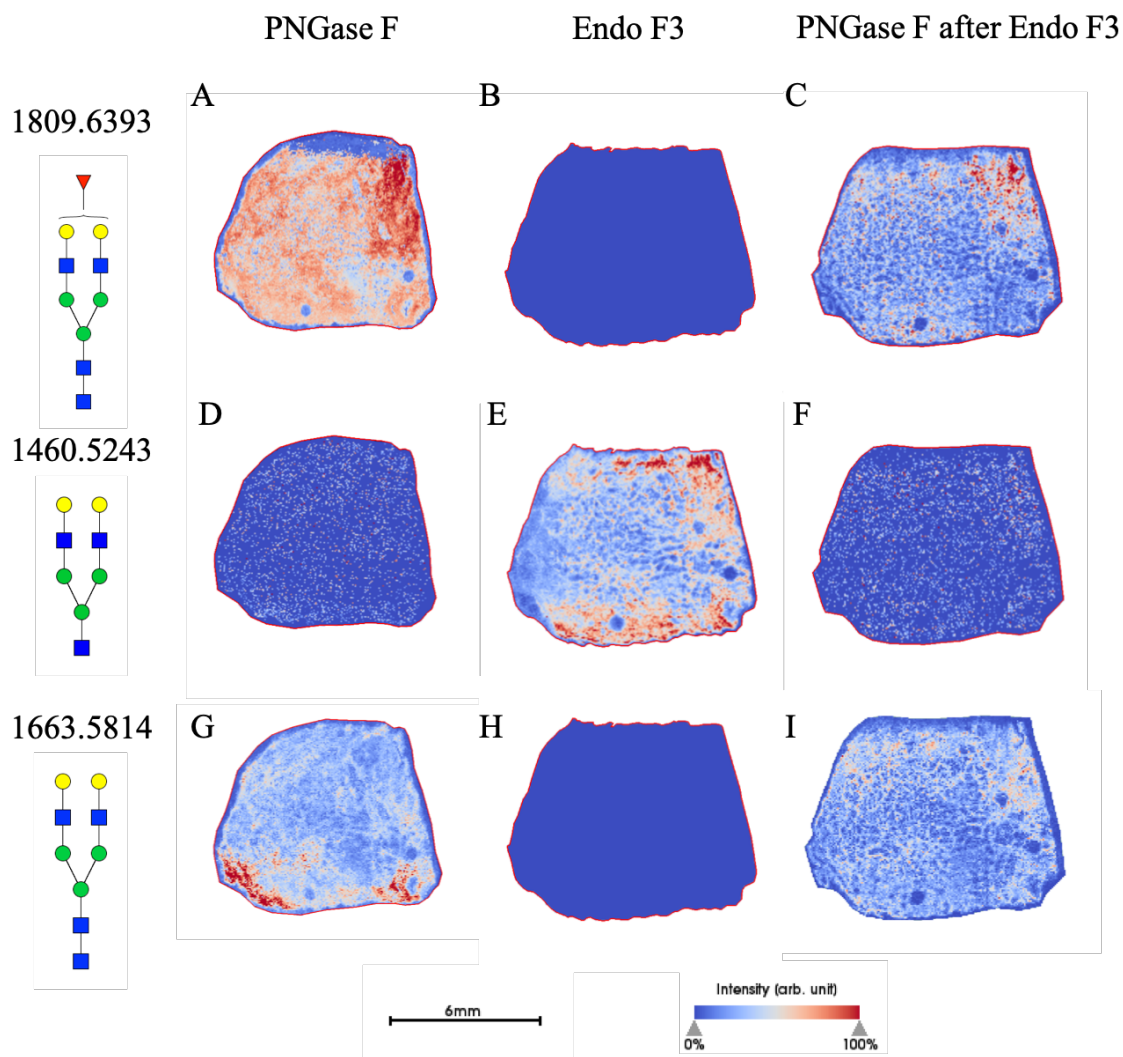
**Figure 23: Generalized Workflow of Endo F3 and PNGase F Treatments.**

Begin by cutting tissue from FFPE block at 5 $\mu$ m on to slide; I) Heat slide at 60°C for one hour; II) Dewax in series of xylene, ethanol dilutions and water, antigen retrieval in citraconic buffer; III) Apply Endo F3 to tissue; IV) Incubate in humidity chamber at 37°C for 2 hours; V) Apply CHCA matrix to tissue; VI) Image on MALDI-FT-ICR; VII) Clear matrix and glycans with ethanol dilutions, high pH and low pH washes; VII) Apply PNGase F to tissue; IX) Incubate in humidity chamber at 37°C for 2 hours; X) Apply CHCA matrix to tissue; XI) Image on MALDI-FT-ICR

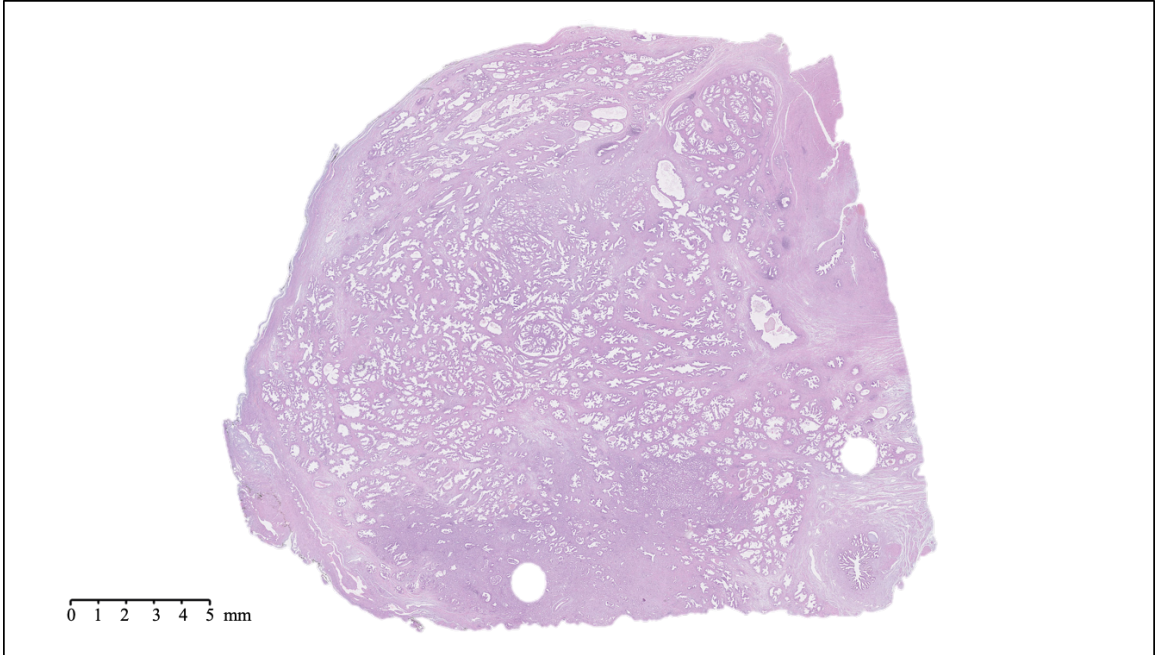
not contain a core fucose residue, similar to what was observed via HPLC (Figure 24). The benefit of tissue imaging is the conservation of spatial localization for the analytes without the need for microdissection or solubilization, and this work maintains this advantage as shown in Figure 25. Following analysis of Endo F3 application on tissue, we found over 30 N-linked glycans to be core fucosylated (Tables 6 and 7) and the main N-glycans found to be core fucosylated are demonstrated in Table 8. These N-glycans also showed localization to specific regions of the tissue. For Figure 25, a prostate cancer tissue section (Figure 26) underwent a variety of treatments where the first column of images represents masses for the tissue following a general PNGase F digestion, the second column represents an Endo F3 digestion, and finally the last column represents an Endo F3 digestion, wash, and sequential PNGase F digestion as described above. As shown in the first row of Figure 25, we see the distribution of the N-glycan A2G2F (1809.6393 predicted m/z) with PNGase F digestion (Figure 25A), a serial tissue section with Endo F3 digestion (Figure 25B) and that same tissue section washed and a sequential PNGase F digestion applied (Figure 25C). These results show that we are not getting any PNGase F cleavage activity on the glycans with our Endo F3 digestion, but still able to achieve the same spatial distribution of the glycans following an Endo F3 digestion, albeit at a lower overall intensity relative to the initial PNGase F digestion. The second row of Figure 25 shows the truncated N-glycan F(6)A2G2 (1460.5023 predicted m/z) following the same treatments. Figure 25D shows that we do not observe this mass following PNGase F digestion, however in 25E, we



**Figure 24. Full Mass Spectra for Prostate Cancer Tissues Treated with Endo F3 and PNGase F.** Further N-glycan identification from the coFull example mass spectra are represented for PNGase F (blue, top) and for Endo F3 (green, bottom) applied prostate cancer tissues. Two major PNGase F fucosylated glycans and their Endo F3 counterparts are highlighted, along with the observed corresponding mass shift. Responding mass spectra peaks can be found in Supplemental Table 1 for PNGase F and Supplemental Table 2 for Endo F3.



**Figure 25: Prostate Cancer Tissues Analyzed with Multiple Enzymes.** Prostate cancer tissue section undergoing PNGase F treatment (A,D,G), Endo F3 treatment (B, E, H) or sequential PNGase F treatment following a wash of the Endo F3 treated tissue (C, F, I). A known core fucosylated glycan, A2G2F distribution is shown for the PNGase F mass of 1809.6393 m/z (A, B, C) and for the Endo F3 treated mass shift of 1460.5023 m/z (D, E, F). Finally, high mannose glycan Man8 (1743.5810 m/z) distribution is shown (G, H, I). Scale bar and intensity bar are included.



**Figure 26: Hematoxylin and Eosin Stain of Prostate Cancer Tissue**

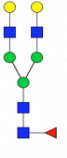
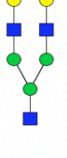
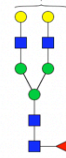
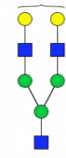
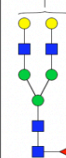
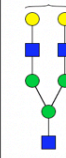
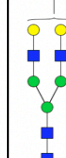
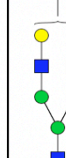
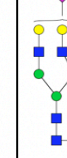



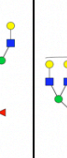

**Table 6: Master List of N-Linked Glycans with PNGase F.** Master list of all N-linked glycans found using PNGase F treatment, along with the corresponding mass error and compositional structure.

Composition	Theoretical m/z	Observed m/z	Error in PPM
Hex5HexNAc2 + 1Na	1257.4226	1257.4337	8.8443
Hex6HexNAc2 + 1Na	1419.4754	1419.4974	15.5276
Hex4dHex1HexNAc3 + 1Na	1444.5071	1444.5368	20.5544
Hex3dHex1HexNAc4 + 1Na	1485.5337	1485.5605	18.0144
Hex4HexNAc4 + 1Na	1501.5286	1501.5592	20.4065
Hex7HexNAc2 + 1Na	1581.5282	1581.5788	31.9634
Hex4dHex1HexNAc4 + 1Na	1647.5865	1647.6356	29.7714
Hex5HexNAc4 + 1Na	1663.5814	1663.6206	23.5883
Hex3dHex1HexNAc5 + 1Na	1688.6131	1688.6569	25.9154
Hex4HexNAc5 + 1Na	1704.6080	1704.6473	23.0792
Hex8HexNAc2 + 1Na	1743.5810	1743.6242	24.7944
Hex5dHex1HexNAc4 + 1Na	1809.6393	1809.7042	35.8806
Hex4dHex1HexNAc5 + 1Na	1850.6659	1850.7334	36.4955
Hex9HexNAc2 + 1Na	1905.6338	1905.7078	38.8432
Hex5HexNAc4NeuAc1 + 2Na	1976.6666	1976.7341	34.1848
Hex5dHex1HexNAc5 + 1Na	2012.7187	2012.7656	23.3222
Hex6HexNAc5 + 1Na	2028.7136	2028.7927	38.9661
Hex5dHex1HexNAc4NeuAc1 + 2Na	2122.7245	2122.7869	29.3924
Hex5dHex2HexNAc5 + 1Na	2158.7766	2158.8272	23.4304
Hex6dHex1HexNAc5 + 1Na	2174.7715	2174.8578	39.6828
Hex4dHex2HexNAc5NeuAc1 + 1Na	2287.8192	2287.8214	0.9577
Hex6HexNAc5NeuAc1 + 1Na	2319.8090	2319.8982	38.4562
Hex6dHex2HexNAc5 + 1Na	2320.8294	2320.9157	37.2026
Hex6HexNAc5NeuAc1 + 2Na	2341.7988	2341.8272	12.1240
Hex6dHex1HexNAc6 + 1Na	2377.8509	2377.8963	19.0975
Hex7HexNAc6 + 1Na	2393.8458	2393.9259	33.4779
Hex9HexNAc3NeuAc1 + 2Na	2421.7984	2421.9534	64.0235
Hex6dHex3HexNAc5 + 1Na	2466.8873	2466.9264	15.8422
Hex6dHex1HexNAc5NeuAc1 + 2Na	2487.8567	2487.9189	25.0022
Hex7dHex1HexNAc6 + 1Na	2539.9037	2539.9544	19.9736
Hex7dHex2HexNAc6 + 1Na	2685.9616	2685.9904	10.7265
Hex8HexNAc7 + 1Na	2758.9780	2759.0200	15.2230
Hex6dHex3HexNAc5NeuAc1 + 2Na	2779.9725	2779.9808	3.0043
Hex7dHex3HexNAc6 + 1Na	2832.0195	2832.0103	3.2376
Hex8dHex1HexNAc7 + 1Na	2905.0359	2905.0798	15.1017

**Table 7: Master List of N-Linked Glycans with Endo F3.** Master list of all N-linked glycans found using Endo F3 treatment, along with the corresponding mass error and compositional structure.

Composition	Parent m/z	Endo F3 Theoretical m/z	Observed m/z	Error in PPM
Hex4HexNAc3 + 1Na	1647.5865	1298.4495	1298.4477	1.4086
Hex3HexNAc4 + 1Na	1688.6131	1339.4761	1339.4328	32.2955
Hex4dHex1HexNAc3 + 1Na	1793.6444	1444.5074	1444.5008	4.5891
Hex5HexNAc3 + 1Na	1809.6393	1460.5023	1460.5117	6.4505
Hex3dHex1HexNAc4 + 1Na	1834.6710	1485.5340	1485.4574	51.5565
Hex4HexNAc4 + 1Na	1850.6659	1501.5289	1501.5338	3.2906
Hex3HexNAc5 + 1Na	1891.6925	1542.5555	1542.5491	4.1678
Hex4dHex1HexNAc4 + 1Na	1996.7238	1647.5868	1647.5796	4.3694
Hex5HexNAc4 + 1Na	2012.7187	1663.5817	1663.6029	12.7262
Hex4HexNAc5 + 1Na	2053.7453	1704.6083	1704.5639	26.0758
Hex5HexNAc3NeuAc1 + 1Na	2100.7347	1751.5977	1751.5814	9.2995
Hex5dHex1HexNAc4 + 1Na	2158.7766	1809.6396	1809.6271	6.9180
Hex6HexNAc4 + 1Na	2174.7715	1825.6345	1825.6245	5.4660
Hex5dHex2HexNAc4 + 1Na	2304.8345	1955.6975	1955.6420	28.3935
Hex6dHex1HexNAc4 + 1Na	2320.8294	1971.6924	1971.6772	7.6934
Hex5HexNAc4NeuAc1 + 2Na	2325.8039	1976.6669	1976.6640	1.4408
Hex5dHex1HexNAc5 + 1Na	2361.8560	2012.7190	2012.7185	0.2579
Hex6HexNAc5 + 1Na	2377.8509	2028.7139	2028.7039	4.9386
Hex5HexNAc3NeuAc2 + 1Na	2391.8301	2042.6931	2042.6896	1.7178
Hex5HexNAc3NeuAc2 + 2Na	2413.8199	2064.6829	2064.6713	5.5883
Hex6HexNAc4NeuAc1 + 1Na	2465.8669	2116.7299	2116.7179	5.6875
Hex6HexNAc4NeuAc1 + 2Na	2487.8567	2138.7197	2138.7127	3.2580
Hex5HexNAc5NeuAc1 + 1Na	2506.8935	2157.7565	2157.7176	18.0368
Hex6dHex1HexNAc5 + 1Na	2523.9088	2174.7718	2174.7546	7.8900
Hex7HexNAc5 + 1Na	2539.9037	2190.7667	2190.8064	18.1128
Hex6dHex2HexNAc5 + 1Na	2669.9667	2320.8297	2320.8643	14.9046
Hex7dHex1HexNAc5 + 1Na	2685.9616	2336.8246	2336.7482	32.6850
Hex6HexNAc5NeuAc1 + 2Na	2690.9361	2341.7991	2341.8213	9.4936
Hex7HexNAc5 + 1Na	2742.9831	2393.8461	2393.8870	17.0942
Hex6dHex1HexNAc5NeuAc1 + 1Na	2815.0042	2465.8672	2465.8742	2.8392
Hex7dHex2HexNAc5 + 1Na	2832.0195	2482.8825	2482.8719	4.2729
Hex7dHex1HexNAc6 + 1Na	2889.0410	2539.9040	2539.9865	32.4937
Hex8HexNAc6 + 1Na	2905.0359	2555.8989	2555.9095	4.1516
Hex7dHex2HexNAc6 + 1Na	3035.0989	2685.9619	2685.9424	7.2745
Hex7dHex1HexNAc6NeuAc1 + 1Na	3180.1364	2830.9994	2831.1097	38.9442
Hex7dHex1HexNAc6NeuAc1 + 2Na	3202.1262	2852.9892	2853.0869	34.2560
Hex9HexNAc7 + 1Na	3270.1681	2921.0311	2921.1289	33.4885

**Table 8: Main Core Fucosylated N-Glycans Found in Prostate Tissue.** Six main N-glycans found to be core fucosylated in prostate tissue, showing the mass to charge ratio, composition, and structure for both the PNGase F and Endo F3 cleaved glycoforms.

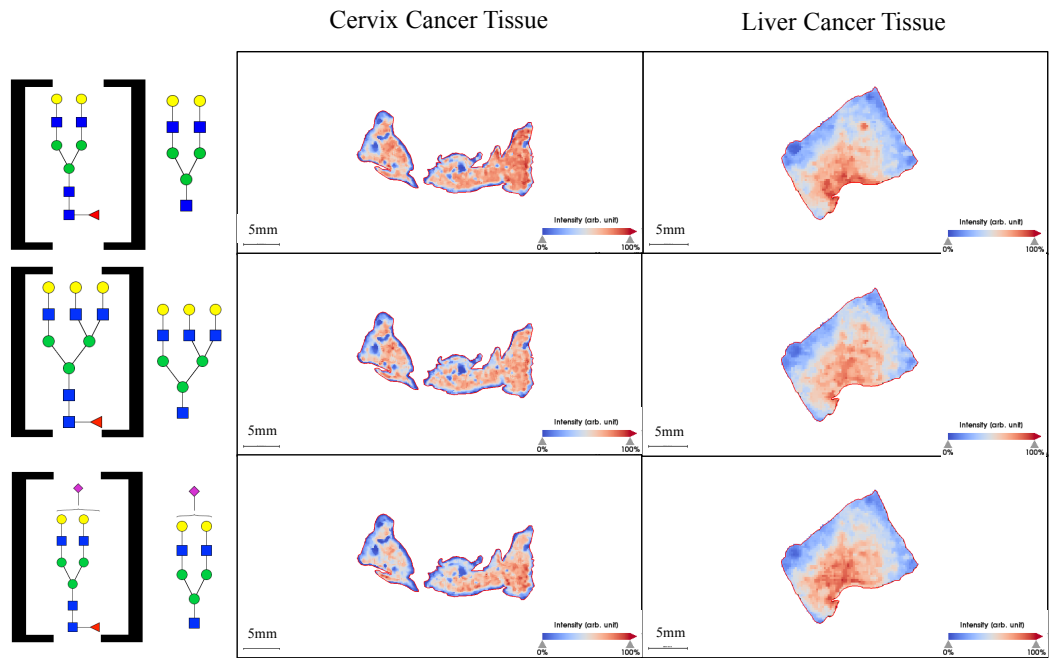
<b>PNGase F Composition</b>	<b>PNGase F Structure</b>	<b>PNGase F m/z</b>	<b>Endo F3 Composition</b>	<b>Endo F3 m/z</b>	<b>Endo F3 Structure</b>
Hex5dHex1HexNAc4 + 1Na		1809.6393	Hex5HexNAc3 + 1Na	1460.5023	
Hex5dHex1HexNAc5 + 1Na		2012.7187	Hex5HexNAc4 + 1Na	1663.5817	
Hex5dHex1HexNAc4 NeuAc1 + 1Na		2100.7347	Hex5HexNAc3 NeuAc1 + 1Na	1751.5977	
Hex6dHex1HexNAc5 + 1Na		2174.7715	Hex6HexNAc4 + 1Na	1825.6345	
Hex6dHex1HexNAc5 NeuAc1 + 1Na		2465.8669	Hex6HexNAc4 NeuAc1 + 1Na	2116.7299	
Hex7dHex1HexNAc6 + 1Na		2539.9037	Hex7HexNAc5 + 1Na	2190.7667	
Hex9dHex1HexNAc8 + 1Na		3270.1681	Hex9HexNAc7 + 1Na	2921.0311	



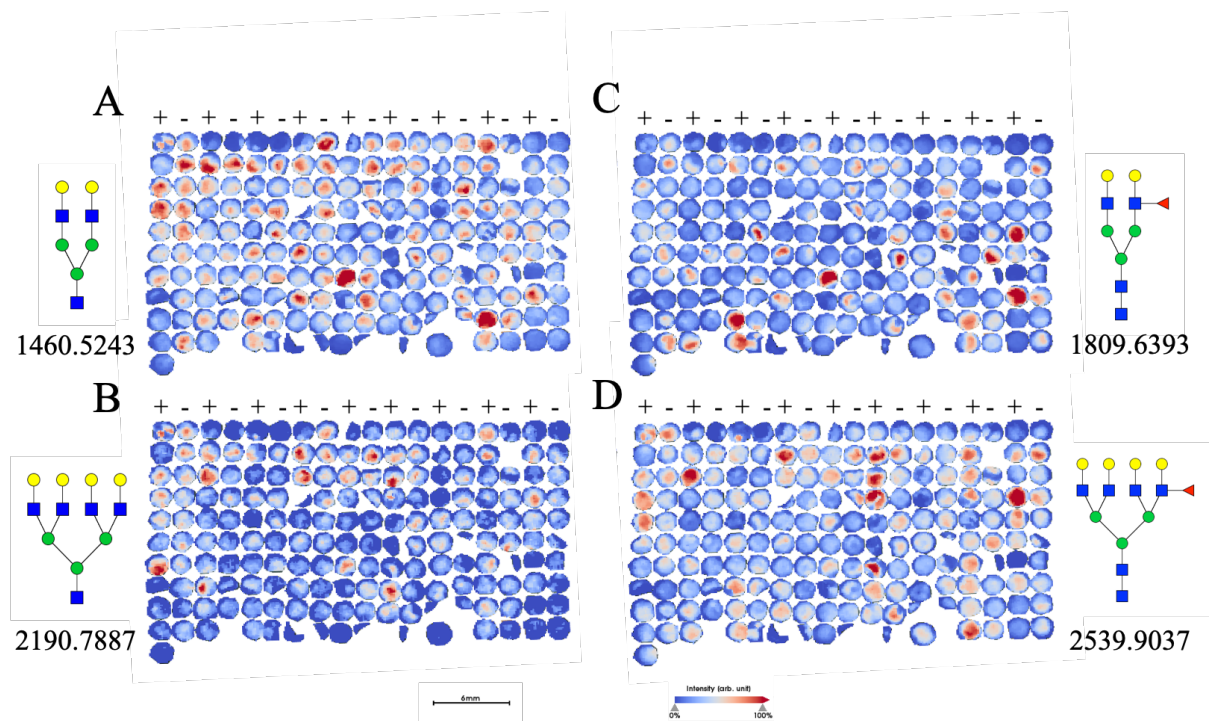
see this mass following the Endo F3 digestion as expected. We are also able to efficiently remove the Endo F3 cleaved glycans following washes and PNGase F application as shown by 25F. Finally, in the third row of Figure 25, we show the distribution of the N-glycan Man8 (1743.5810 predicted m/z) which should never contain a core fucose. Again, we see a similar situation to the first row, with PNGase F cleavage (Figure 25G), no cleavage with Endo F3 (Figure 25H), and a less efficient salvage with a sequential PNGase F digestion (Figure 25I). To be certain that the effectiveness of the Endo F3 digestion was not tissue specific, we also performed similar digestions on multiple tissue types (Figure 27).

#### **4.4.3. Endo F3 Application to Patient Tumor Microarray**

With the ability to determine core versus outer arm fucosylation, we then wanted to apply this technique to patient samples to determine the relevancy of this technique for determining clinically relevant factors. As previously described, core fucose is implicated in many cancer progressions, so we applied the Endo F3 followed by PNGase F protocol to a purchased hepatocellular carcinoma TMA set (US Biomax) as previously analyzed by our group [161]. In Figure 28, we see two different fucosylated glycans implicated in the paper, A2G2F and A4G4F (1809.6393 and 2539.9037 predicted m/z respectively). Figures 28A and 28B represent F(6)A2G2 and F(6)A4G4 in their reduced forms following Endo F3 digestion (1460.5023 and 2190.7667 predicted m/z respectively) while 28C and 28D represent the sequential wash and PNGase F digestion for non-core fucosylated A2G2F and A4G4F. In examining the results, we see that there are



**Figure 27: Multiple Tissue Types Treated with Endo F3.** Cervical cancer tissue as well as hepatocellular carcinoma tissue were treated with Endo F3 and analyzed for core fucosylated glycans shown to the left with the parent glycan structures shown in the brackets.

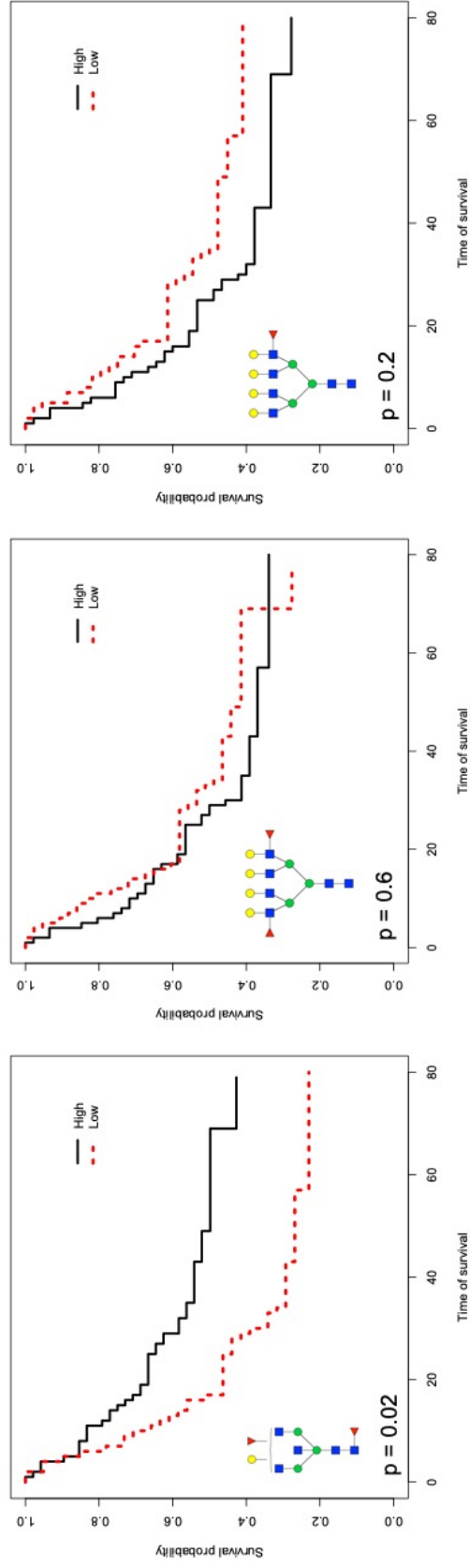


**Figure 28: Patient TMA Treated with Multiple Enzymatic Digestions.** Patient TMAs treated with Endo F3 (A and B) followed with a PNGase F digestion (C and D) are shown with two prominent core fucosylated glycans of F(6)A2G2 and F(6)A4G4 abundance shown (1809.6393 and 2539.9037 m/z respectively). (+) indicates cancerous tissue while (-) indicates normal, untransformed tissue.

TMA cores that contain relatively more of the core fucosylated versions of the glycans while some contain relatively more non-core fucosylated glycans. While this is not an absolute quantitation and more direct analysis will be required to determine the abundance of core vs. outer arm fucosylation, this work shows promise that we can further parse out the underlying mechanisms and difference resulting from the tumor heterogeneity between patients.

#### **4.4.4. Core Fucosylation and Patient Survival Outcomes**

Because this study utilized the same patient tumor microarrays studied previously, I was afforded numerous benefits, specifically in terms of direct comparison to previous studies, as well maintaining the patient survival characteristics. In previously published results involving this tumor microarray, it was shown that survival probability is decreased in patients with elevated levels of A4G4F (m/z 2539.904) and A4G4F2 (m/z 2685.968) [161], however when examined under the dual-enzymatic conditions described above, elevated levels of the glycan above involving no core fucosylation shows no significant difference in survival probability (Figure 29, middle and right panel), thus demonstrating the effectiveness and relevance of the dual-enzymatic approach. In contrast, the left panel of Figure 29 demonstrates a bisecting, core fucosylated N-glycan (A3G1F2, 1996.724 m/z) that was never previously implicated in having a significant effect on patient survival. When examined via Endo F3, however, we see that there is a significant difference in survival probability for patients above the median level as compared to those below the median value. This N-glycan is



**Figure 29: Survival Plots for Patient TMA Using Endo F3 and PNGase F.** The first panel demonstrates a bisected N-glycan shown to be significantly related to lower survival outcomes when upregulated, an N-glycan not identified in previous studies. In the middle and right panel, previously identified glycans shown to have an impact on survival are shown to have no significant impact on survival when determined to be non-core fucosylated.

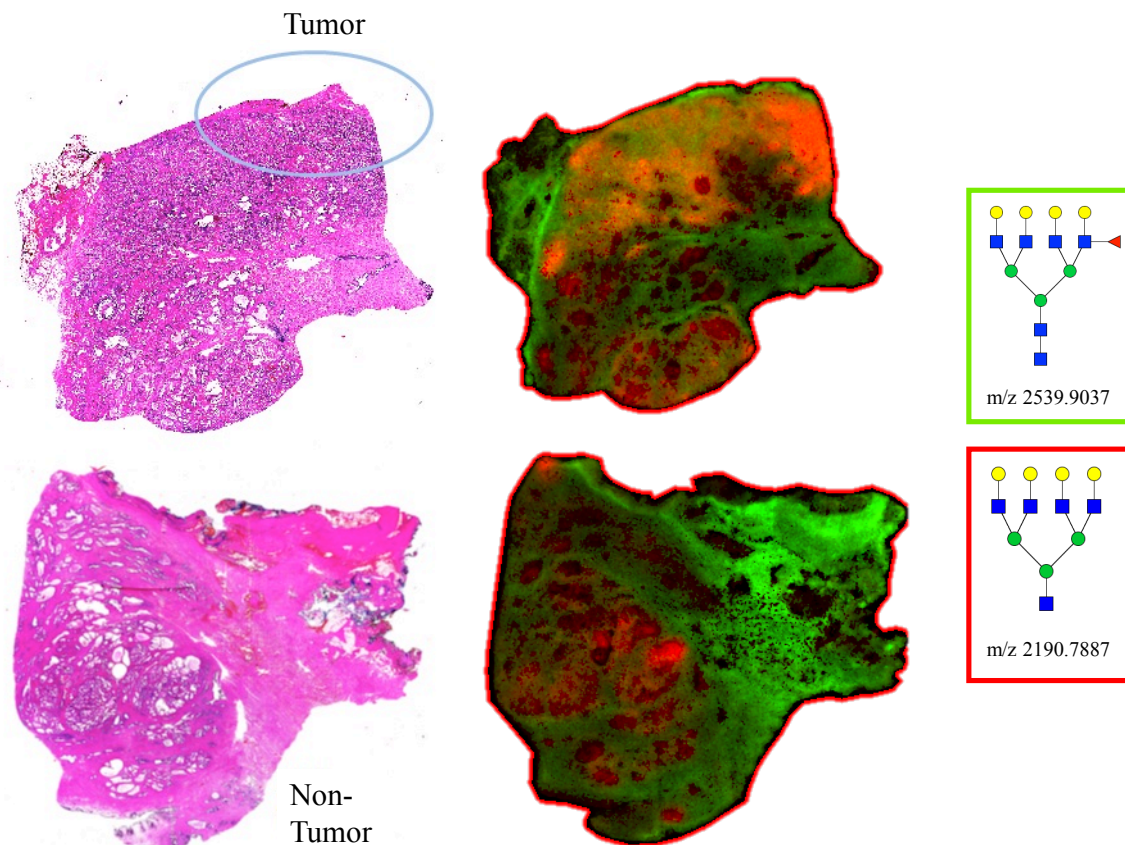
a prime example of those that may not be apparently obvious when looking at all glycans together, but when further classifying the structures of these glycans can reveal new potential differentiating structures. Along with this, further studies are needed and ongoing regarding patient outcomes and tumor grading and scoring in terms of core versus outer arm fucosylation.

#### **4.4.5. Combined Application of PNGase F and Endo F3**

To further explore the process of the dual enzyme cleavage, initial experiments were conducted to determine the possibility of mixing both PNGase F and Endo F3 in one spray. The enzymes were initially combined at a 3:1, 1:1 or 1:3 ratio of Endo F3 and PNGase F and it was found that the lower concentration of Endo F3 was better suited for cleaving both core and non-core fucosylated N-glycans (data not shown). From this, the enzyme concentration of Endo F3 was lowered further, and experiments were done to show a 1:20 ratio of Endo F3 to PNGase. This demonstrated the best spectra regarding efficient cleavage of all N-glycans of interest (Figure 30). Efficiency and control experiments are still needed; however, this is a promising start to further optimizing the dual enzymatic workflow.

#### **4.5. Discussion**

As we know, fucosylation of N-linked glycans has been associated with several types of cancer [294], especially changes in the addition of core  $\alpha$ -1,6-linked fucose is associated with the development of hepatocellular carcinoma



**Figure 30: Mixture of PNGase F and Endo F3.** Representative images of Endo F3 and PNGase F mixed at an approximate ratio of 1:20 respectively (middle). Hematoxylin and eosin stain included (left), as well as structure and m/z (right) for distribution of the green N-glycan (m/z 2539.9037) and red N-glycan (m/z 2190.7887)

(HCC) [295]. Comparing to PNGase F, Endo F3 is working more efficiently and selectively on core  $\alpha$ -1,6-linked fucosylated structures. Without the interference and noises of all the other complex glycans released by PNGase F, we can focus on the core  $\alpha$ -1,6-linked fucosylated structures when comparing patients' samples to healthy controls. This was demonstrated above in Figure 25, showing that comparing core versus outer arm fucosylation does vary patient to patient, though the underlying mechanism is still unclear.

The most notable benefit of this work is the addressing of one major drawback to the applied methods for MALDI imaging mass spectrometry. With the ability to distinguish between the anomeric linkages of the fucose additions of the glycans, more in-depth analysis of tissue is capable without the use of serial sections or other structural elucidation techniques that lose the spatial localization afforded with imaging, such as proteomic analysis or ion mobility [296].

Additionally, this methodology has the potential to improve glycopeptide analysis in the field of proteomics. With the residual HexNAc and Fucose residue left following the Endo F3 cleavage, this could be utilized in proteomic analysis as a more specific precursor ion. Used appropriately, this precursor ion could be indicative of glycopeptides that contained core fucosylated N-glycans, further elucidating the structural motifs of the attached N-glycans with well-established and easy to perform proteomic analyses, such as electron-transfer dissociation (ETD).



While this technique can effectively determine core fucosylated N-glycans, the protocol will still require further optimization to reach efficiency levels similar to that of the PNGase F. As it stands now, the PNGase F digestion is working on N-glycans substantially more efficiently than the Endo F3 digestion, rendering quantitative analysis difficult. However, despite the flaw in quantitative analysis, the qualitative abilities of the data are able to further elucidate the localization and relative abundance of these core fucosylated glycans. With this information, more distinct patterns and features can be acquired from the tissue image, allowing for more comprehensive analysis of tissue imaging and glycosylation as it relates to tumor heterogeneity.

# **Chapter 5: Determination of Sialic Acid Linkage via MALDI- IMS**

## 5.1. Introduction

As referenced previously in chapter three, one significant challenge in examining the changes in glycosylation via MALDI imaging mass spectrometry is the inherent misrepresentation of sialylated N-glycan species. Compared to other glycosidic bonds, sialic acids are among the least stable, with loss and decay very common following the ionization step in a MALDI imaging mass spectrometry experiment [297]. Further, the presence of a strong negative charge on sialic acids increases ionization inefficiencies, particularly when experiments are conducted in positive ion mode, as most N-glycan experiments are [297]. Because of these challenges presented, in order to accurately quantitate changes in sialylated N-glycans for biomarker analysis, sialic acids on N-glycans must be stabilized for true representation of sialylated species in the human N-glycome. For example, the loss of one sialic acid on a bi-antennary glycan would result in an apparent increase in the non-sialylated bi-antennary glycan, thus implying a lower population of sialylated species and a miscalculated increase in the unmodified bi-antennary glycan. This being said, qualitative analysis is still passable, with increases in more complex glycans being associated with the tumor regardless of sialylation state, allowing for differentiation of tumor versus non-tumor via N-glycan analysis. However, the true identity of these glycans will continue to remain elusive without determining the accurate sialylation state on these higher branched structures. With this being the case, many groups have strived to further stabilize these sialic acids on

N-glycans and differentiate the specific glycosidic linkages of the sialic acid on the N-glycan.

In this chapter, I will discuss the application of two different chemical derivatization techniques that were adapted from previously published articles. The first method that will be discussed is ethyl esterification (EE), adapted from Reiding *et. al.* in 2014 [241]. This technique creates an EE reaction on  $\alpha$ 2,6 linked sialic acids while performing lactonization on the  $\alpha$ 2,3 sialic acids, allowing for stabilization of the sialic acid to the galactose and differentiating the two linkage varieties by change in mass. The second method is a double amidation reaction (AA) performed on tissue, adapted from Holst *et. al.* in 2016 [262]. While utilizing many of the same materials, this method varies slightly from the EE method through substituting ethanol for dimethylamine, resulting an initial dimethylamidation reaction, followed by a secondary amidation reaction via ammonia in water. In  $\alpha$ 2,6 sialic acids, this results in the formation of an amide group, while in  $\alpha$ 2,3 sialic acids, the lactone is reopened by the second reaction and an amine group is added. Again, this method allows for stabilization of sialic acids to prevent pre- and post-source ionization decay and allows for elucidation of the linkage of the sialic acid through the resulting mass shift. In this chapter, I will discuss the initial experiments using these two methods, the rationale behind choosing one method over the other, the application to human liver tissue, and potential applications of the methodology.

## **5.2. Materials and Methods**

### **5.2.1. Tissue Samples and Relevant Materials**

Formalin-fixed paraffin-embedded (FFPE) tissue blocks of whole tissue (specifically prostate and liver) were obtained from the Medical University of South Carolina Hollings Cancer Center Biorepository and Tissue Analysis Shared Resource. A set of 12 tissue microarray (TMA) FFPE slides containing varying patient data were obtained from collaborators in Heidelberg, Germany.

For the chemical derivatization, 1-Hydroxybenzotriazole hydrate (HOBt) with dimethylamine and approximately 20 percent water was purchased from Sigma-Aldrich (St. Louis, MO). 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (EDC) was purchased from Alfa Aesar (Haverhill, MA) and stored in a sealed bag with drierite and sealed with parafilm due to its moisture and air sensitivity. Finally, dimethyl sulfoxide (DMSO) and 28 to 30 percent ammonia in water was purchased from Fisher Scientific (Hampton, NH). Additionally, cover glass was purchased from ThermoFisher Scientific (Waltham, MA) to secure solutions to tissue.

### **5.2.2. Washes and Rehydration**

Washes and dehydration steps were performed similarly to what was described in Section 3.3.3.

### **5.2.3. Ethyl Esterification Chemical Derivatization**

Ethyl esterification was adapted from Reiding *et. al.* (2014) for direct tissue analysis. Briefly, immediately following the deparaffinization and rehydration washes, a mixture of EDC and HOBt was made for on-tissue derivatization. To create the EE solution of 0.25 M EDC and 0.25 M HOBt, 0.675g of HOBt was combined with 0.959g of EDC, 1mL of HPLC H<sub>2</sub>O, and 20mL of 200-proof ethanol. A glass mailer was then filled with the EE solution and the slide fully immersed. The lid was then parafilmmed shut to prevent any evaporation of the solution and the glass mailer placed into a 37°C oven for one hour. Finally, the slide was carefully removed to avoid tissue disruption as much as possible and a series of washes were used to remove the reagents. These washes consisted of two minutes in 200-proof EtOH twice, 10 minutes in Carnoy's Solution (60 percent ethanol, 30 percent chloroform, and 10 percent acetic acid) twice, two minutes in HPLC H<sub>2</sub>O, two minutes in 200-proof EtOH twice, and a rinse of 1 percent trifluoroacetic acid (TFA) in 200-proof EtOH for 30 seconds. It is important to note that following each wash step, the solutions were replaced to remove the reagents as completely as possible. The slide was then taken directly, without the chance to dry, for further tissue processing using the experimental MALDI imaging workflow.

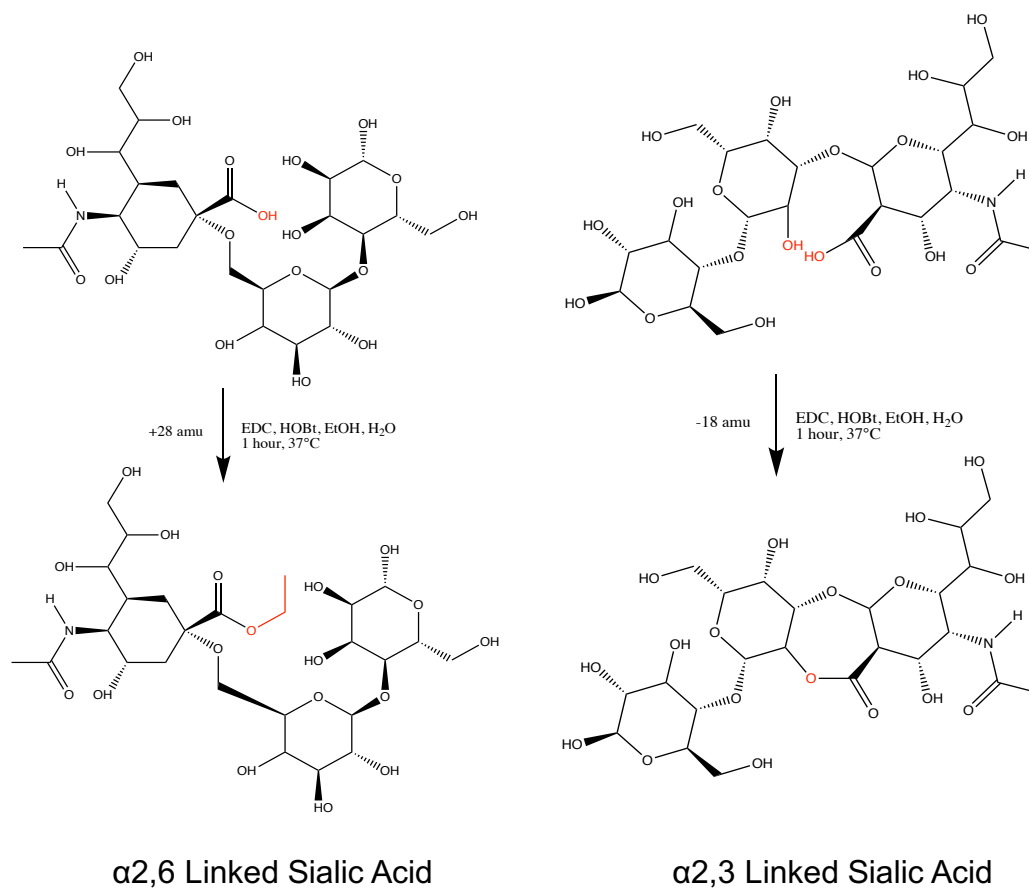
#### **5.2.3.1. Reaction Schematic**

As stated above, the EE method utilizes EDC, HOBt, and ethanol to derivatize sialic acids through either the esterification on  $\alpha$ 2,6 linked sialic acids

or the removal of water from  $\alpha$ 2,3 sialic acids, both of which cause a discernable shift in m/z in a MALDI imaging experiment. In  $\alpha$ 2,6 sialic acids, the exposed carboxylic acid reacts with the ethanol through the action of the EDC and HOBt, resulting in an esterification reaction and the mass shift of approximately 28.05 a.m.u. (+46.07 for the addition of the ethanol and -18.02 for the loss of water). EE acts as a protective group for the sialic acid, allowing for less in-source decay. Conversely,  $\alpha$ 2,3 linked sialic acids perform a lactonization event with the neighboring galactose, where the exposed carboxylic acid is joined with an exposed hydroxyl group, creating a more stable glycosidic bond and a mass shift of -18.02 a.m.u. from the loss of water in the lactone formation. Figure 31 provides a more in-depth look at the structural changes associated with EE and highlights the differences based on sialic acid linkages.

#### **5.2.4. Amidation-Amidation Chemical Derivatization**

Adapted from Holst *et. al.* in 2016, the AA reaction was performed as described with three slight modifications: 1) the reaction volume was reduced to 200 $\mu$ L, 2) a washing step was added between the first and second amidation, and 3) the amount of HOBt was adjusted for 20 percent water composition. Briefly, reaction solution 1 was created by combining 22 $\mu$ L of EDC, 42.2mg of HOBt, 15.8 $\mu$ L of dimethylamine, and 0.5mL of DMSO to create a 0.25/0.5/0.25 M solution of EDC, HOBt, and dimethylamine, respectively. 200 $\mu$ L of reaction solution 1 was then applied to the tissue using a pipette tip, careful to completely cover the tissue without disruption. The tissue was then covered with a glass



**Figure 31: Ethyl Esterification Derivatization of Sialic Acids.** Chemical structures of  $\alpha 2,3$  and  $\alpha 2,6$  sialic acids are shown above, as well as the modification made during the Ethyl Esterification chemical derivatization process. Changes are highlighted in red, and the conditions for the derivatization and associated mass shift are shown as well.



coverslip and placed into either a pyrex petri dish with parafilm or a pyrex container with airtight sealing capabilities. The container was then placed inside a 60°C oven for 1 hour.

Following incubation, the slides were then removed and the coverslip was gently removed, careful to not disrupt the tissue adherence to the slide. The slide was then placed perpendicularly on a paper towel to remove the majority of the reaction solution. To remove the rest, a vacuum flask apparatus was set up to allow for suctioning of remaining liquid in a more precise manner. Using a pipette tip and with as close contact as possible without touching the tissue, the remainder of the visible liquid was aspirated, though it is important to note that over-drying the tissue was avoided as it causes reagents to adhere more to the tissue. 200µL of DMSO was applied to the tissue and then aspirated in a similar fashion as described above, and this process was repeated a total of two more times.

Next, the second reaction was prepared with 350µL of DMSO and 150µL of 28 to 30 percent ammonia in water. The second reaction solution was then applied in a similar fashion to the first, with all tissue covered and a coverslip placed over the tissue to seal. The slide was then placed inside the pyrex chamber to protect from evaporation and placed in the 60°C oven for 2 hours. Washing of the slide was performed exactly as described above. Finally, the slide was then rinsed in a series of washes as described in section 5.2.3. (ethanol, Carnoy's, water, and TFA). The slides then proceeded directly to further processing without being allowed to dry.

#### **5.2.4.1. Reaction Schematic 1**

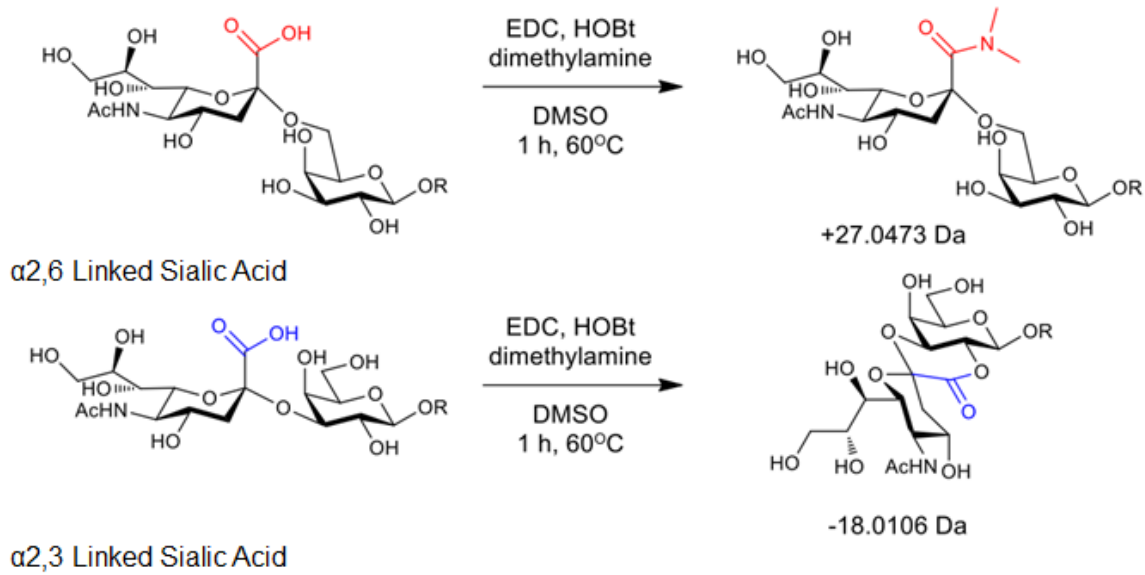
In contrast to the EE reaction, the AA reaction utilizes dimethylamine instead of ethanol. In examining  $\alpha$ 2,3 sialic acids, the reaction mechanism is unchanged, creating a lactone between the sialic acid and neighboring galactose and resulting in a loss of water. However, in  $\alpha$ 2,6 linked sialic acids, the carboxylic acid reacts with the dimethylamine, creating a stable dimethylamide group and a loss of water, resulting in a mass shift of approximately +27.05 a.m.u. Figure 32 highlights the structural changes of both sialic acid and the change in mass for each.

#### **5.2.4.2. Reaction Schematic 2**

Unlike EE, the AA reaction has a second reaction step, resulting in amidation of the  $\alpha$ 2,3 linked sialic acids as well. In this step, the lactone formed from the first reaction is hydrolyzed and the ammonia creates a stable amide group on the newly formed carboxylic acid. This results in a total mass difference of -0.984 a.m.u. for  $\alpha$ 2,3 linked sialic acids following both reactions. The  $\alpha$ 2,6 linked sialic acids remain completely unchanged throughout the duration of the second reaction. Figure 33 details the conditions and structural changes for the  $\alpha$ 2,3 linked sialic acids and the resulting mass differences.

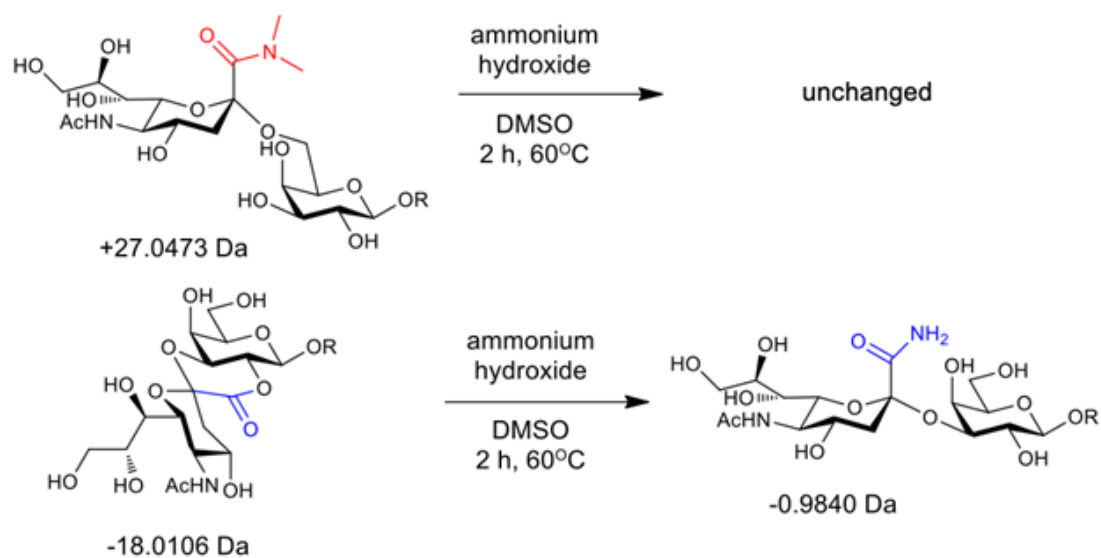
#### **5.2.5. Tissue Preparation and N-Glycan Imaging**

For the tissues that underwent the EE chemical derivatization, they were then treated as previously described in sections 3.3.3. to 3.3.5., proceeding



**Figure 32: Reaction Schematic for Amidation-Amidation Reaction #1.**

Schematic of sialic acid derivatization via amidation-amidation reaction. Top: reaction scheme for  $\alpha$ 2,6 linked sialic acids; Bottom: reaction scheme for  $\alpha$ 2,3 linked sialic acids. Shown for addition of EDC, HOBt, and dimethylamine. Changes in  $\alpha$ 2,6 sialic acids are highlighted in red while changes in  $\alpha$ 2,3 sialic acids are highlighted in blue.



**Figure 33: Reaction Schematic for Amidation-Amidation Reaction #2.**

Schematic of sialic acid derivatization via amidation-amidation reaction. Top: reaction scheme for  $\alpha 2,6$  linked sialic acids; Bottom: reaction scheme for  $\alpha 2,3$  linked sialic acids. Shown for addition of ammonium hydroxide to further distinguish  $\alpha 2,3$  linked sialic acids. Changes in  $\alpha 2,6$  sialic acids are highlighted in red while changes in  $\alpha 2,3$  sialic acids are highlighted in blue.

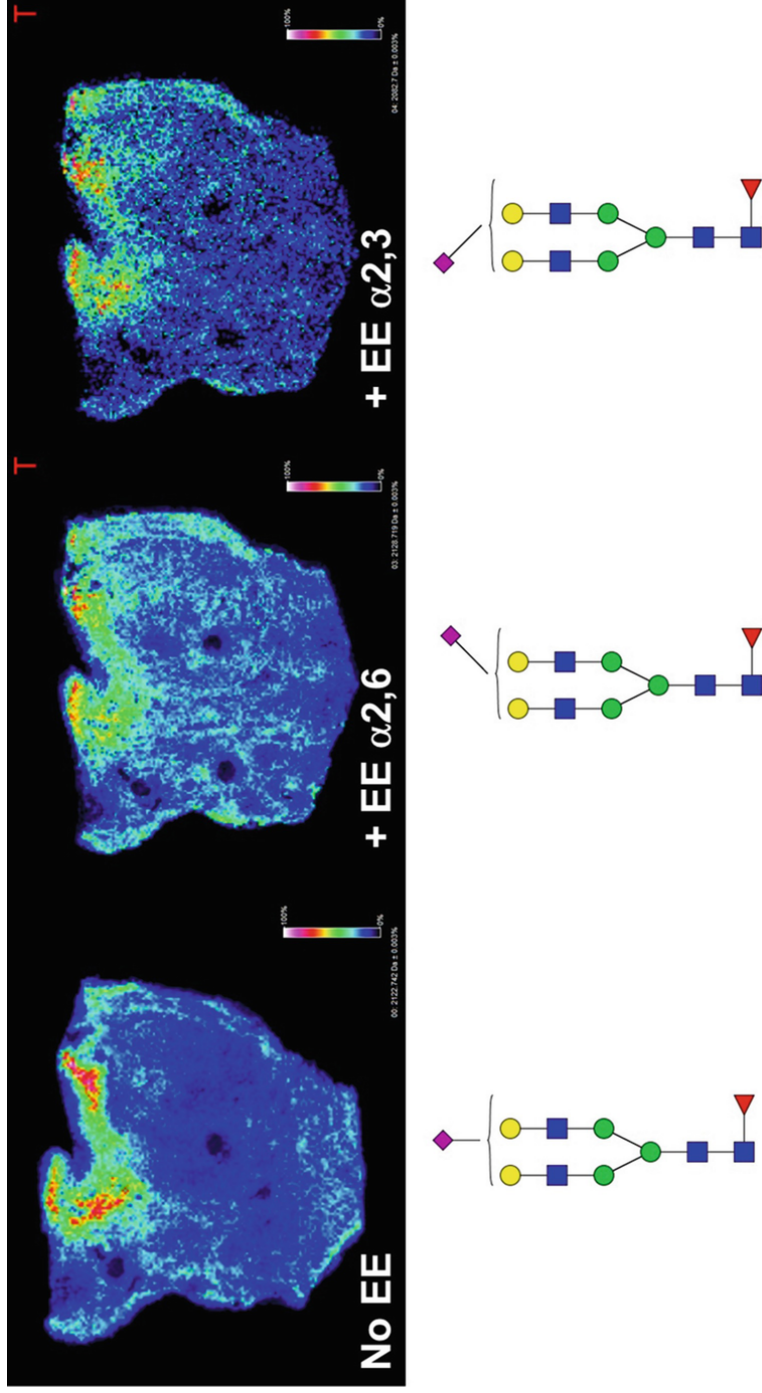
through an antigen retrieval step, enzymatic digestion and matrix deposition, and N-glycan imaging on both the MALDI FT-ICR and MALDI-TOF systems [259, 267].

For tissues that underwent the AA reaction, a standard decloaker was used to proceed with antigen retrieval. Briefly, the overall basin and slide holder jars were filled to the appropriate level, and a mailer containing the slides was filled with approximately 20 mL of citraconic buffer (as described in section 3.3.3.) to completely fill the mailer and one snap of the lid was left open to prevent bursting of the mailer and to allow steam to exit. The decloaker parameters were set to 115°C for 15 minutes with a pressure setting of 5.4 psi. The slides were placed in the decloaker for the preheating and allowed to run the full cycle. Finally, once the cycle was finished and the decloaker had cooled to 95°C, the mailer was then cooled and washed similar to the process described in section 3.3.3. The reason for the shift in method was due to two factors: 1) the vegetable steamer previously used was no longer functional and 2) other members of the lab performed analysis on decloaker settings compared to vegetable steamer and rice cooker settings and found that these settings were optimal for tissues with high fat content or for tissues that had undergone a chemical stabilization (data not shown). The rest of the tissue processing occurred similarly to sections 3.3.4. and 3.3.5.

## 5.3. Results

### 5.3.1. Ethyl Esterification On-tissue

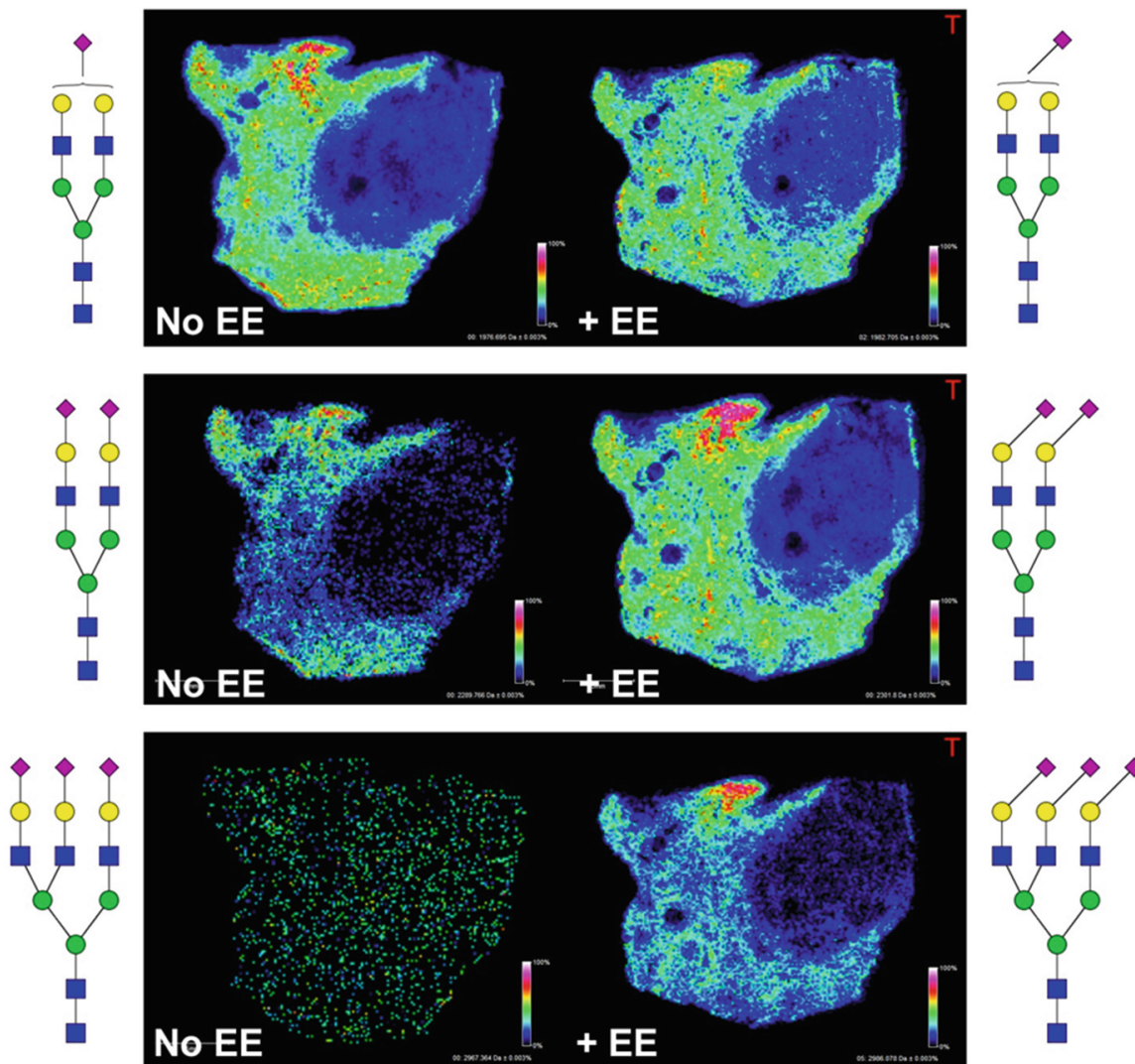
Initially, the method for EE was created for released glycans, so the method was adapted for on-tissue application with the adjustments being made by myself in collaboration with Dr. Peggi Angel for the washing procedure. Once the protocol was adapted and applied to tissue, the first test was to determine the efficiency of stabilization for sialic acids post-imaging analysis. In Figure 34, initial experiments were done to show the efficiency of sialic acid stabilization and differentiation, mainly looking to examine the overall retention of sialylated species and the associated mass shifts corresponding to the linkage of the sialic acid. An example glycan that is mono-sialylated, mono-fucosylated, and bi-antennary (2100.7347 m/z, A2G2F1S1) is shown without EE, the chemical derivatization for an  $\alpha$ 2,3 sialic acid, and for an  $\alpha$ 2,6 sialic acid. The localization of the sialylated species is clearly visible in the prostate FFPE tissue, demonstrating the preference to smooth muscle and collagen [76]; however, when looking at the chemically derived masses, there is a slight variation to the localization. Not only do we see a higher representation of the N-glycan along the sides of the tissue, we see representatively similar amounts of sialylated species following the chemical derivatization, supporting the claim that sialylated species are indeed more stabilized following the protocol, as well as differentiated by mass as to be expected (approximately +28 a.m.u. for  $\alpha$ 2,6 and -18 a.m.u. for  $\alpha$ 2,3).



**Figure 34: Stabilization of Sialic Acids by On-Tissue Ethyl Esterification.** Ethyl esterification (EE) was adapted from a published protocol (Reiding et al. 2014) for a prostate cancer FFPE tissue and analyzed by MALDI-FT-ICR IMS. Shown are the examples for a mono-sialylated core fucose bi-antennary glycan, without EE (left panel) or with EE for  $\alpha 2,6$  (+28 amu) (middle panel) or  $\alpha 2,3$  (-18 amu) (right panel) linkages

While this was a promising start for derivatizing on tissue, these initial experiments were performed on the MALDI FT-ICR platform, which is known to image sialylated species more efficiently than a time-of-flight platform. Therefore, it was imperative to better prove the stabilization aspect by examining higher sialylated species than a mono-sialylated N-glycan. As mentioned previously, higher sialylated species may lose some sialic acid modifications while retaining others, resulting in a misrepresentation of sialylated N-glycans. So, to determine the efficiency of on-tissue stabilization, three N-glycans were examined with varying amounts of sialic acid modifications (A2G2S1, 1954.6768 m/z; A2G2S2, 2245.7722 m/z; and A3G3S3, 2923.9896 m/z). Figure 35 shows the same prostate FFPE tissue that underwent both EE and no modification to the tissue. When examining these three N-glycans, the stabilization is clearly visible when comparing the left half of the images with the right. On the left, as the number of sialic acid modifications increase, a clear loss of signal is shown, demonstrating reduced abundance of these sialylated N-glycan species, particularly the tri-sialylated N-glycan. In contrast, the right half of the images show a greater retention in signal of these multi-sialylated glycans, even showing a clear localization pattern for the tri-sialylated N-glycan that was not even discernable without the chemical modification. Therefore, we were able to clearly show a stabilization effect stemming from this N-glycan derivatization on tissue as compared to untreated tissue, as well as determine specific N-glycan linkages based on the corresponding mass shift post-treatment.





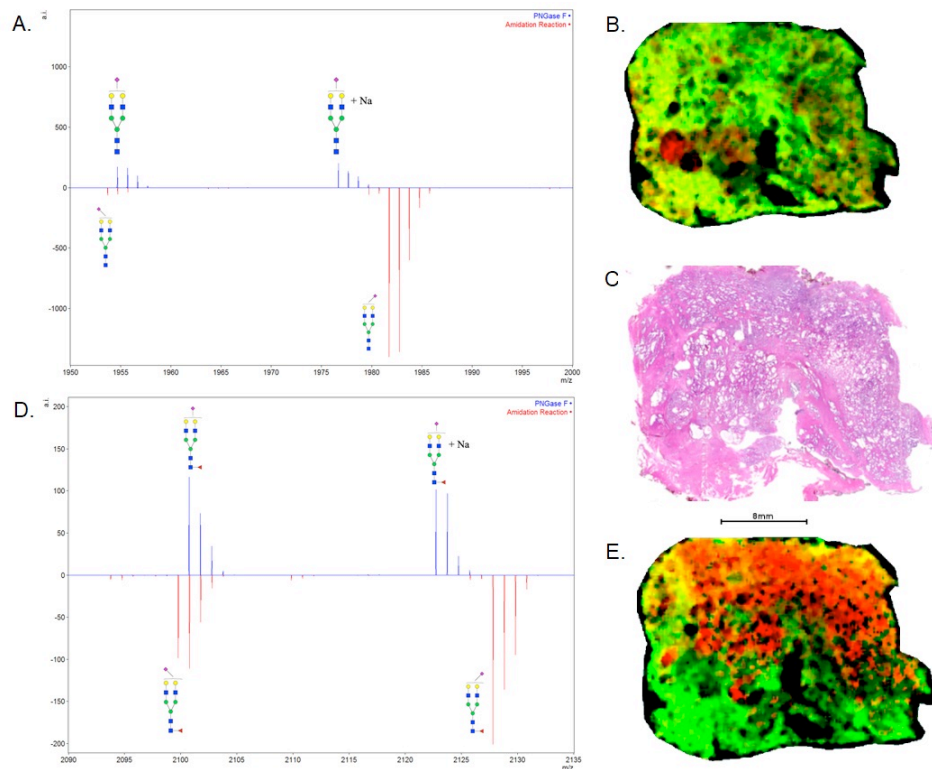
**Figure 35: Stabilization of Multi-Sialylated Tissue Glycans by EE.** MALDI-FT-ICR IMS data from the same ethyl esterified prostate tissue is shown for mono-, di-, and tri-sialylated bi-antennary glycans. Images in each panel on the left side represent non-ethyl esterified glycans, and the right side images are after EE. Also shown for the EE examples are the  $\alpha$ -2,6 sialic acid linkage structures

### 5.3.2. Amidation-Amidation On-tissue

In addition to applying the ethylation reaction on tissue, myself and another member of the lab worked toward adapting the method from Holst *et. al.* (2016) for on-tissue analysis. Using a combination of the paper adaptation and slight adjustments made from the ethylation reaction protocol, a working amidation procedure was once again validated on FFPE prostate tissue. Table 9 demonstrates example changes in parent N-glycan masses according to sialic acid linkages, while Figure 36 demonstrates the N-glycan amidation spectra and the resulting images from prostate FFPE tissue section. Figure 36A shows a non-amidated spectra on top in blue and an amidated spectra on the bottom in red for the mono-sialylated bi-antennary (A2G2S1) N-glycan (1954.6768 m/z). Here, it is clearly shown that the amidation results in a higher overall signal, implying stabilization of the sialylated N-glycan. Figure 36B is an overlay of the  $\alpha$ 2,3 linked N-glycan in red and the  $\alpha$ 2,6 linked N-glycan in green, demonstrating the implications in localization we can see in FFPE tissue through differentiating the linkage specificities of the same N-glycan. Figure 36D shows a similar spectrum as to 36A, but for the mono-sialylated, mono-fucosylated, bi-antennary (A2G2S1F1) N-glycan (2100.7347 m/z). Again, we can see an increase in intensity for the peaks of the amidated tissue, implying stabilization. Finally, we see the overlay of the  $\alpha$ 2,3 linked N-glycan in red and the  $\alpha$ 2,6 linked N-glycan in green in Figure 36E. This distribution is localized completely differently from 36B, showing how differentiating these glycans even further can dictate trends in understanding tissue N-glycan localization in terms of disease state, tissue morphology, etc.

**Table 9: Main Sialylated N-Glycans and Mass Shift from AA Reaction.** Shown above are frequent examples of sialylated N-glycans, their compositional structure, parent m/z, and the changes in m/z based on linkage following the amidation-amidation chemical derivatization.

N-Glycan Composition	Predicted Structure	Parent m/z	One $\alpha$ 2,3 linked sialic acid m/z	One $\alpha$ 2,6 linked sialic acid m/z	Two $\alpha$ 2,3 linked sialic acid m/z	Two $\alpha$ 2,6 linked sialic acid m/z	Two mixed sialic acid linkages m/z
Hex5HexNAc4NeuAc1 + 1Na		1954.6768	1953.6928	1981.7241	--	--	--
Hex5dHex1HexNAc4 NeuAc1 + 1Na		2100.7347	2099.7507	2127.7820	--	--	--
Hex6HexNAc5NeuAc1 + 1Na		2319.8090	2318.8250	2346.8563	--	--	--
Hex5HexNAc4NeuAc2 + 1Na		2245.7722	--	--	2243.8042	2299.8668	2271.8355
Hex5dHex1HexNAc4 NeuAc2 + 1Na		2391.8301	--	--	2389.8621	2445.9247	2417.8934

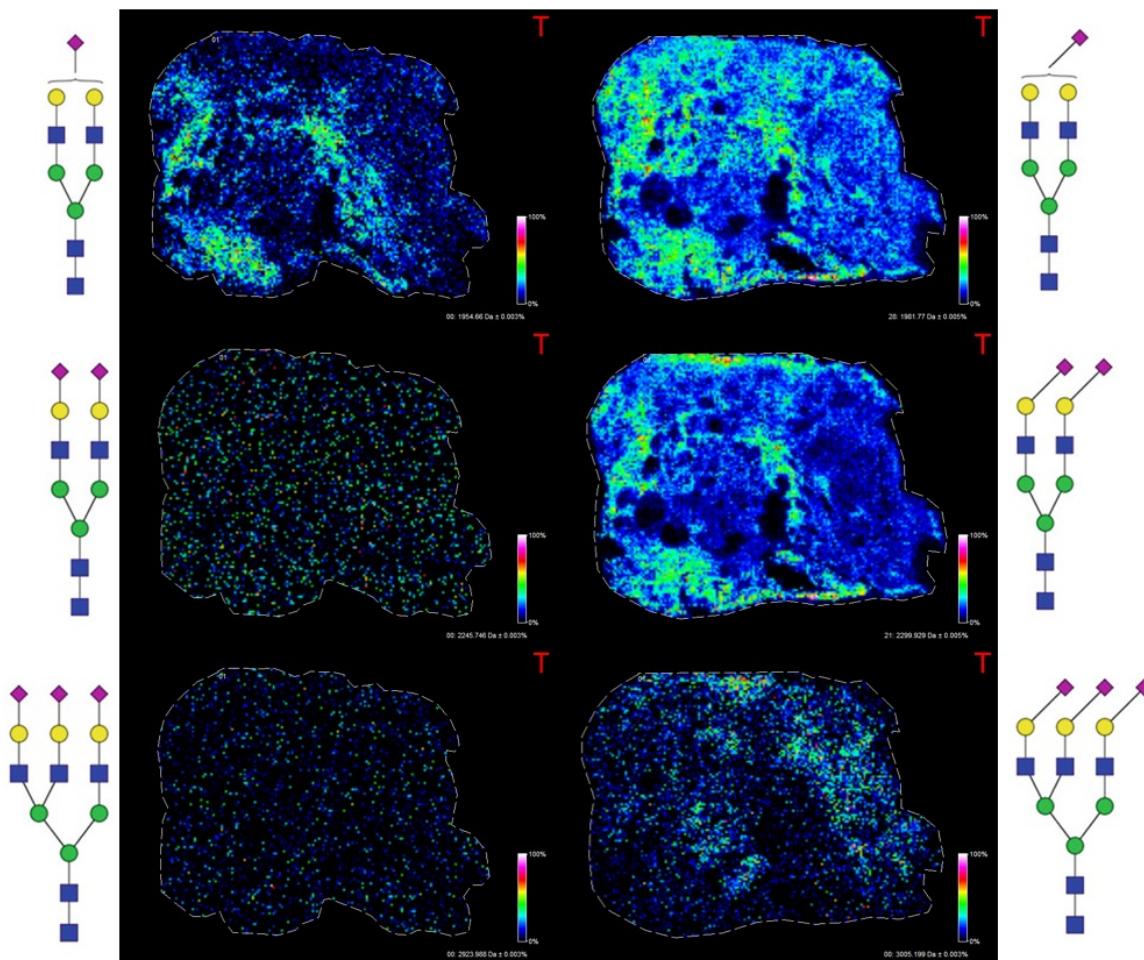


**Figure 36: Amidation-Amidation Chemical Derivatization On-Tissue.** Amidation-amidation reaction done on prostate FFPE tissue and the resulting spectra and images following analysis via MALDI imaging mass spectrometry. A) Spectra for N-glycan A2G2S with and without amidation-amidation. The top spectra in blue represents the non-derivatized N-glycan, with one or two sodium ions. The bottom spectra in red represents the mass shift resulting from the chemical derivatization. Sialic acids angled to the left indicate  $\alpha$ 2,3 linkage while angled to the right represents  $\alpha$ 2,6 linkage. B) Resulting image overlay showing localization of the  $\alpha$ 2,3 linked N-glycan,  $m/z = 1953.741$  (red, -1 m.u.), and  $\alpha$ 2,6 linked,  $1981.782$  (green, +27 m.u.). C) Hematoxylin and eosin stain of the tissue section analyzed. D) Spectra for the fucosylated N-glycan A2G2S1F1 with and without the amidation-amidation reaction. Top spectra in blue is non-derivatized and the bottom spectra in red is showing the mass shift following derivatization. E) Resulting image overlay showing localization of  $\alpha$ 2,3 linked N-glycan,  $m/z = 2099.810$  (red, -1 m.u.) and  $\alpha$ 2,6 linked,  $m/z = 2127.832$  (green, =27 m.u.).

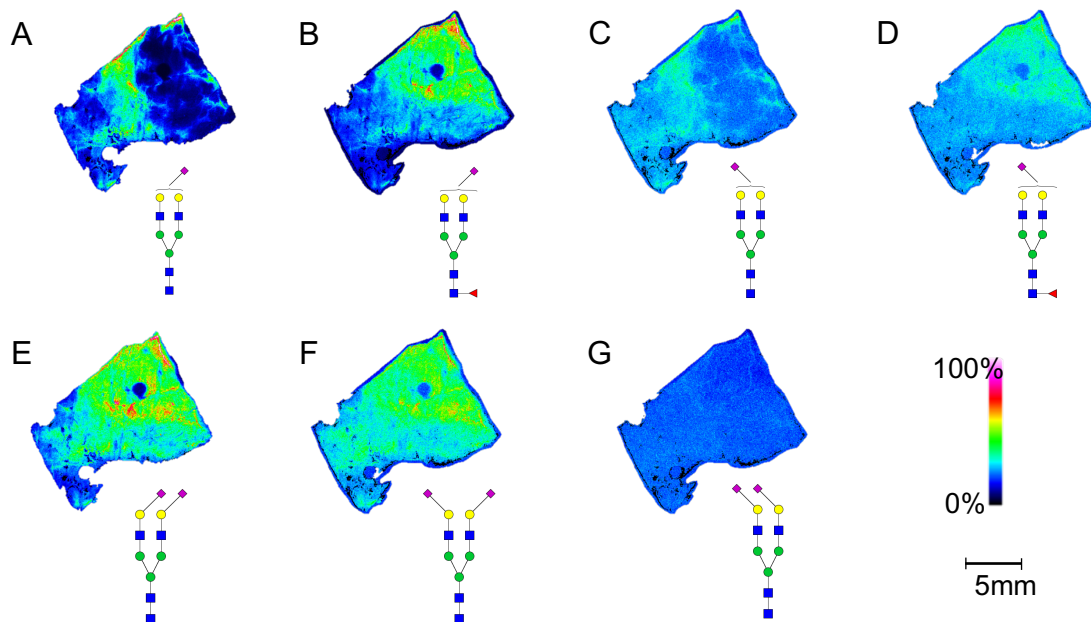
Similar to the EE validation, it was also important to further confirm that this method will increase sialic acid stability by examining multi-sialylated N-glycan structures. Figure 37 is set up comparably to Figure 35 and examines N-glycans with one, two, and three sialic acids before and after amidation. The left column of images represents the non-amidated tissue and a drop in signal intensity and clarity is clearly visible in the di- and tri-sialylated N-glycans. However, when looking at the amidated tissue, we see a continuation of signal intensity when comparing the mono-sialylated N-glycan to the di-sialylated, and even the tri-sialylated shows a distinct localization pattern and higher intensity as compared to the non-amidated tissue. These studies validated the use of amidation to not only differentiate sialic acid linkage, but to also stabilize multi-sialylated N-glycan species, allowing for more accurate qualitative and quantitative N-glycan analysis. Therefore, this methodology was then used to examine sialic acid difference in HCC tissue.

### **5.3.3. Amidation-Amidation on FFPE Human Liver Tissue**

Amidation derivatization was then applied to HCC tissues to further characterize the N-glycome of HCC, in addition to the fucosylation characterization demonstrated in chapter four. Figure 38 shows the amidation protocol applied to the same FFPE HCC tissue that was used in Figure 27 in chapter four. A and C demonstrate the mono-sialylated A2G2S1 N-glycan while B and D demonstrate the mono-sialylated and mono-fucosylated A2G2S1F1 N-glycan. A and B both show the  $\alpha$ 2,6 linkage for the sialic acid and C and D



**Figure 37: Stabilization of Multi-Sialylated Tissue Glycans by AA.** MALDI-FT-ICR IMS data from the same amidated prostate tissue is shown for mono-, di-, and tri-sialylated bi-antennary glycans. Images in each panel on the left side represent non-amidated glycans, and the right side images are after AA. Also shown for the AA examples are the  $\alpha$ -2,6 sialic acid linkage structures

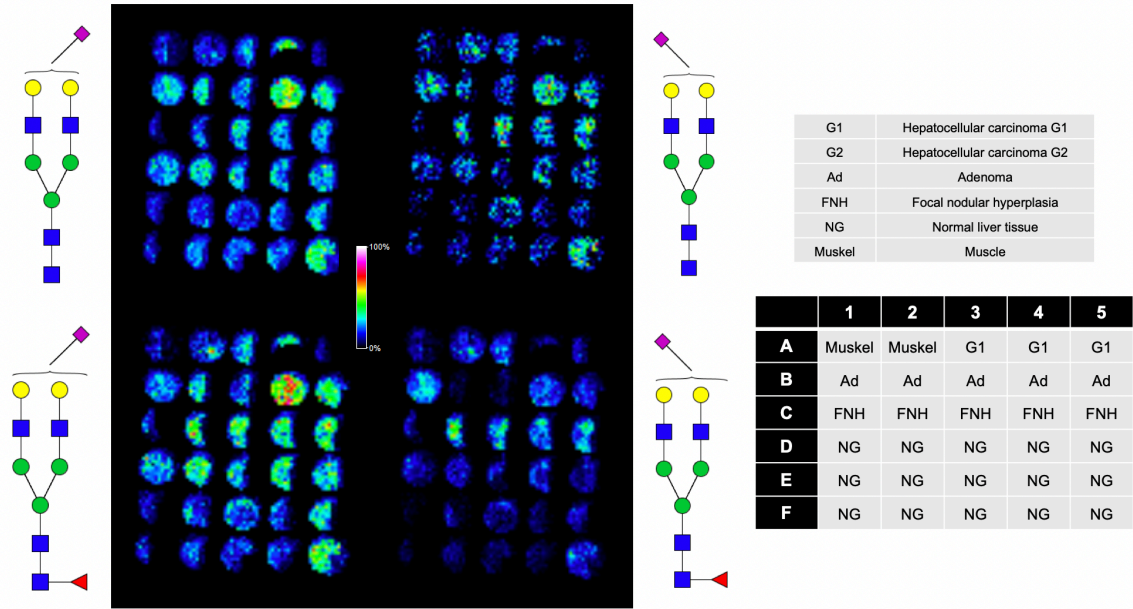


**Figure 38: AA Chemical Derivatization of HCC Tissue.** Following optimization of the amidation-amidation conditions, the derivatization was then applied to representative hepatocellular carcinoma tissues, seen above. The top row shows two different sialylated N-glycan species, one with and without a core fucose, with the  $\alpha$ 2,6 linked sialic acid species in the left two images and the  $\alpha$ 2,3 linked sialic acid species in the right two images. The bottom row shows a di-sialylated N-glycan with the left demonstrating 2  $\alpha$ 2,6 linked sialic acids, the middle representing a mix of both, and the right showing 2  $\alpha$ 2,3 linked sialic acids.

demonstrate the  $\alpha$ 2,3 linkage for the sialic acid. When comparing the  $\alpha$ 2,6 to the  $\alpha$ 2,3 in the tissue, the overall intensity is decreased for the  $\alpha$ 2,3 specific N-glycan species; however, localization is not changed by sialic acid linkage, but rather whether there was a fucose residue present on the structure. Additionally, when examining the bi-sialylated N-glycan shown in E, F, and G, the presence of one or two  $\alpha$ 2,6 sialic acids results in an intense signal localized to the tumor region of the tissue; however, the presence of two  $\alpha$ 2,3 sialic acids is not seen anywhere in the tissue, driving the importance of linkage specificity and the possible importance of  $\alpha$ 2,6 linked sialic acids in HCC.

In conjunction with single tissue imaging applications, the amidation protocol was also applied to TMAs to examine multiple tissues at once to further characterize the changes in terms of tissue type, morphology, or by patient. Unfortunately, the TMA utilized in chapters three and four was not available for amidation. However, as previously mentioned, a collaborator in Heidelberg, Germany sent us a set of twelve TMAs of varying tissue types, pathologies, and number for which to examine the N-glycome. With this set, the amidation derivatization was applied to a few of these TMAs to examine the implications of sialic acid linkages on a broader scale. In Figure 39, the key to the TMA and the tissue types is shown to the right, with representative images of the same mono-sialylated N-glycans as shown previously, with and without the fucose residue. Comparable to what was seen in Figure 38, the  $\alpha$ 2,6 linked sialic acid structures were much more intense than the  $\alpha$ 2,3 sialic acid structures, regardless of tissue morphology, pathology, or patient, as shown on the left side of Figure 39. Further





**Figure 39: AA Chemical Derivatization of Human Liver TMA.** Amidation-amidation was applied to a human liver tissue microarray with varying tissue types (see keys to the right) and the resulting changes in sialylation are shown to the right. In general,  $\alpha$ 2,6 sialylation was much more prominent than  $\alpha$ 2,3 sialic acid regardless of additional modifications or tissue morphology type.

studies are planned to continue analysis of these TMA sets to further distinguish the differentiation of fucosylation or sialylation in HCC tissues, as well as examine these trends at a patient-by-patient basis as was done in chapter three.

#### **5.4. Discussion**

As briefly touched upon in the introduction, the need to further characterize sialylation is backed by ever-growing developments in the role sialylation plays in disease states, particularly in cancer development, progression, invasion, and metastasis. As reviewed by many groups, alterations in sialic acids are consistent in many aspects of cancer biology, such as reduced adhesion of tumor cells to the ECM to promote invasion and metastasis, the masking of innate immunity pathways in the complement system to promote tumor survival, and alterations in immune cell receptors and selectins [298-300]. Further, hyper-sialylation and sialic acid linkage specificity has been implicated in several specific cancers, such as ovarian cancer, prostate, oral cancer, and even HCC [77, 163-166, 301-304]. These implications were the main driver of the work in this chapter, specifically in examining those sialic acid changes in HCC.

On the broadest scale, total and free sialic acid has been examined in liver disease. In looking at the total sialic acid, the different etiologies of liver disease were not significantly different, but compared to normal controls, the total sialic acid concentration was decreased remarkably. In contrast, the free sialic acid concentrations varied between etiologies, showing that nonalcoholic cirrhotic liver serum contained less free sialic acid than those of a toxic etiology [165]. The total

sialic acid concentration has also been shown to differentiate cholangiocarcinoma from HCC with a sensitivity of 82.6 percent and a sensitivity of 83.1 percent [166].

In a more specific view of HCC, sialyltransferase regulation, mucin expression, Thomsen-Friedenreich (TF/Tn) antigen expression, and linkage specification has been examined for the disease state. Previously, the ST6Gal family of sialyltransferases were shown to be upregulated in HCC but not cirrhosis, implying that the alteration of  $\alpha$ 2,6 sialylation of liver glycoproteins happens during neoplastic transformation and can be useful in identifying potential biomarkers for early stage HCC [305]. Additionally, the ST6Gal family has been shown to modulate chemotherapeutic responses for HCC cells, reducing the efficacy of docetaxel treatment in-vivo. These results imply that that this sialyltransferase may play a role in maintaining cancerous cell survival in HCC, consistent with previous reports of upregulation of ST6Gal-I and increased prevalence of  $\alpha$ 2,6 sialic acids [164]. Finally, other groups have examined the expression of mucins and TF/Tn related antigens, where MUC1, MUC2, TF, Tn, sialosyl-Tn, and  $\alpha$ 2,6 linked sialic acids were not readily expressed in normal hepatocyte cells, while HCC showed upregulation of all but MUC2 after comparing expression between normal and HCC tissue to confirm that the results were not from an incomplete glycosylation event [163]. This and all the previous findings are consistent with what we have seen in our liver cancer tissues, with an upregulation of total  $\alpha$ 2,6 sialic acid apparent as opposed to the  $\alpha$ 2,3 conformation, leading to a promising start in further characterizing HCC's total N-

glycome and adding more information to further distinguish tumor from non-tumor in HCC (especially in combination with additional fucosylation studies).

To further discuss the rationale behind choosing the AA chemical derivatization over the EE, the decision lies in the continued benefits of an amide addition as opposed to an ethyl-ester group. Research by Dr. Vivian Lu in the Drake lab has shown that replacement of the second amidation from ammonia to an propargyl amine or azide depending on the linkages results in a mass shift as well as an added benefit with this modification. With these propargyl amine or azide group attachments to the terminal sialic acids, this opens the door for potential “click chemistry” reactions, allowing for bioconjugation of additional molecules to the sialic acids [306]. For example, ongoing work by Dr. Lu has demonstrated effective conjugation of markers such as GFP to the sialic acids on-tissue, as well as the conjugation to magnetic beads off tissue. This conjugation to the sialic acids provides an enhanced benefit when looking at proteomics, allowing for a more targeted approach to the glycoprotein carriers of these modified sialic acids, further optimizing the enrichment of specific N-glycan modifications and their protein carriers. This work is still being conducted with a manuscript in process demonstrating the effectiveness of these clickable sialic acid targets.

# **Chapter 6: Enhanced Glycoproteomic Analysis of Liver Cancer Tissues**

## 6.1. Introduction

While defining the N-glycome in human disease, specifically HCC, is important, the clinical relevance of these modifications to glycosylation is integrally linked with the glycoprotein on which the modification is occurring. Therefore, it is imperative to not only determine the modification of interest, but also correctly identify the protein of interest. Intact glycoproteomics, i.e. simultaneous analysis of a glycopeptide singly or in mixtures, has increased in popularity, with multiple techniques developed within the last decade [172, 213, 215, 272, 307-314]. However, there are inherent difficulties of intact glycoprotein analysis that have made this progress slow-moving. For example, the complexity of an intact glycopeptide structure, the low abundance of glycopeptides in respect to the total protein concentration, and the low ionization efficiency of these glycopeptides all provide challenges in the intact analysis of these glycoproteins [315]. However, as discussed in chapter one, many techniques have recently been developed to combat these challenges, such as CID, HCD, ETD and a combination of these methods [213, 219, 226, 312, 316].

Early studies of glycoproteomics utilized collision-induced dissociation, a technique where ions are accelerated and collide with a neutral molecule (such as helium, nitrogen, or argon) to release their kinetic energy internally and fragment the molecule. CID uses a lower collisional energy, ideally suited for preferential glycan fragmentation while leaving the peptide backbone largely unmodified [317]. One drawback, however, is the low  $m/z$  cutoff, making the identification of larger glycan structures more difficult. Similar to CID, HCD

effectively fragments the sugar structure while also fragmenting a portion of the peptide backbone, producing B- and Y-type ions for improved peptide identification [308]. Unlike CID, however, HCD does not suffer from the low m/z cutoff, making this technique better suited for more complex and higher mass glycans. It is also able to be performed in a stepwise function, changing the collisional energy and obtaining different information over the varying range of collisional energies [310]. While this works for characterization of the glycan moiety, it is ineffective in terms of completely identifying the peptide of interest, as well as locating the site of glycosylation on the peptide. Conversely, ETD is most effective for examining the protein or peptide backbone while leaving the post-translational modifications intact. Briefly, ETD causes fragmentation of large, positively charged molecules through the transfer of an electron, resulting in cleavage to the C- and Z-type ions, providing an advantage for longer, intact peptides and proteins [310]. ETD provides a unique approach when combined with other collisional techniques, allowing for a combination of glycan identification and site-specific localization on associated glycopeptides. Unfortunately, while many experiments to date have utilized these techniques for protein glycosylation studies, very little has been reported in terms of disease or human health implications.

As stated above, one of the many challenges of these techniques is the low abundance of these glycopeptides following tryptic digest, which introduces difficulties in detection and quantification. Therefore, many studies include an enrichment approach of some kind to enhance glycopeptide abundance in the

sample. One enrichment technique commonly used for proteomics is lectin affinity chromatography, where a specific lectin is used to bind glycoproteins carrying the lectin-target glycan of interest initially, thus reducing the pool of potential proteins. For example, *Aleuria aurantia* Lectin (AAL) is commonly used to enrich for fucosylated glycans, further narrowing the proteomic hits of a sample [318]. While these are extremely effective, they often require the homogenization of a sample, so they are largely ineffective when looking at region-specific tissue samples for proteomics.

The following work outlines the combination of imaging mass spectrometry techniques with glycoproteomic approaches to enhance glycopeptide identification for specific N-glycan changes. More specifically, this new methodology combines the previously discussed enzymatic derivatization of core fucosylation on tissue for imaging purposes to enhance detection and characterization of core fucosylated glycopeptides in cancer tissues. With the use of the unique GlcNAc-Fuc tag left on specific core fucosylated peptides, the normally complex glycopeptide spectra are more easily interpreted and reduce ionization complications of larger and more complex core fucosylated N-glycans.

## **6.2. Materials and Methods**

### **6.2.1. Tissue Samples and Relevant Materials**

FFPE liver cancer tissues were obtained from the Medical University of South Carolina Hollings Cancer Center Biorepository and Tissue Analysis Shared Resource. Endoglycosidase F3 was obtained, expressed, and purified by



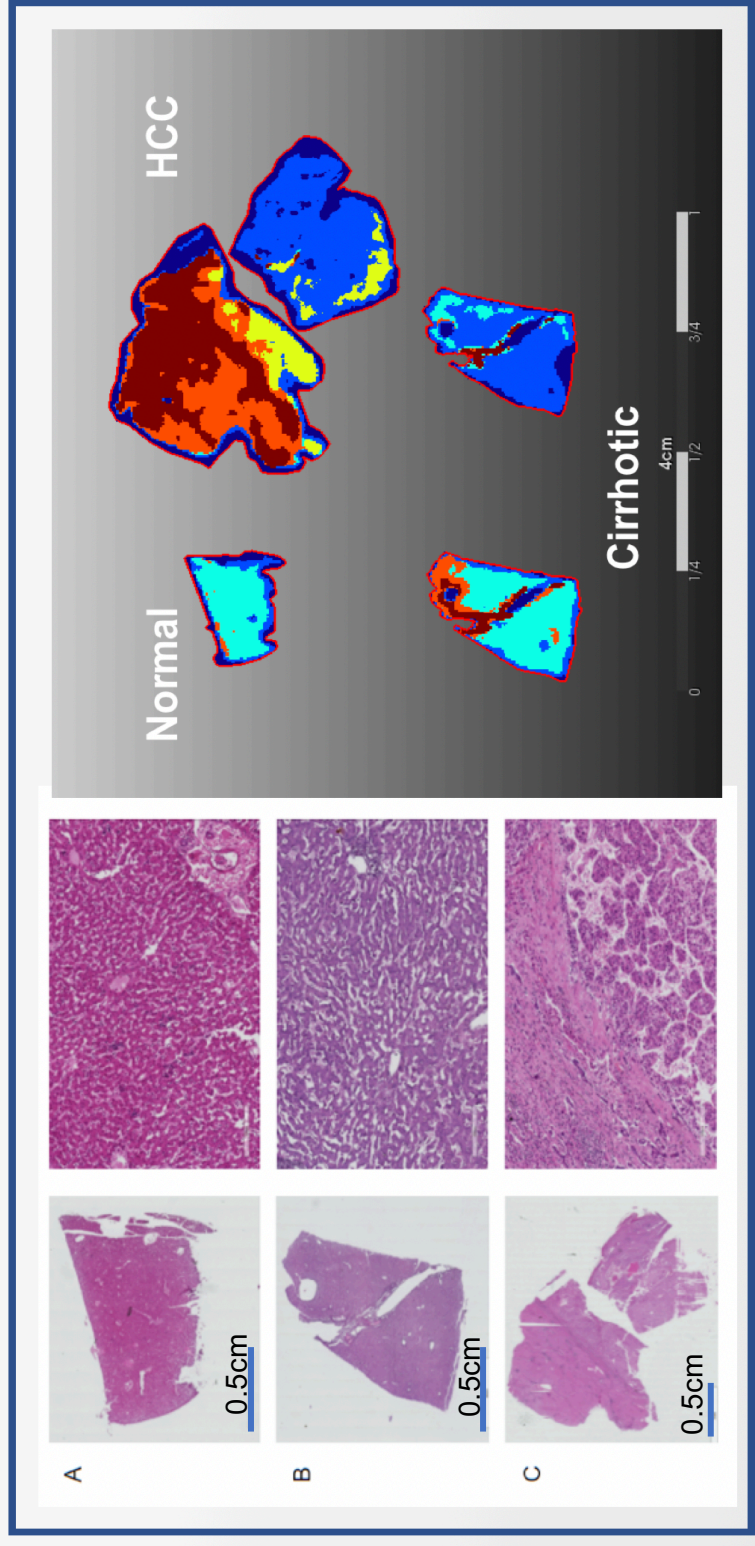
the Mehta laboratory. Pierce Microplate BCA Protein Assay Kit and C18 ZipTip cartridges were purchased from ThermoFisher Scientific (Waltham, MA). Trypsin was purchased from Promega (Madison, WI) at five bottles of 20 $\mu$ g of trypsin. All other solutions and solvents made are generic varieties. For our purposes, solvents were purchased from Fisher Scientific (Hampton, NH).

### **6.2.2. Imaging Characterized N-Glycans**

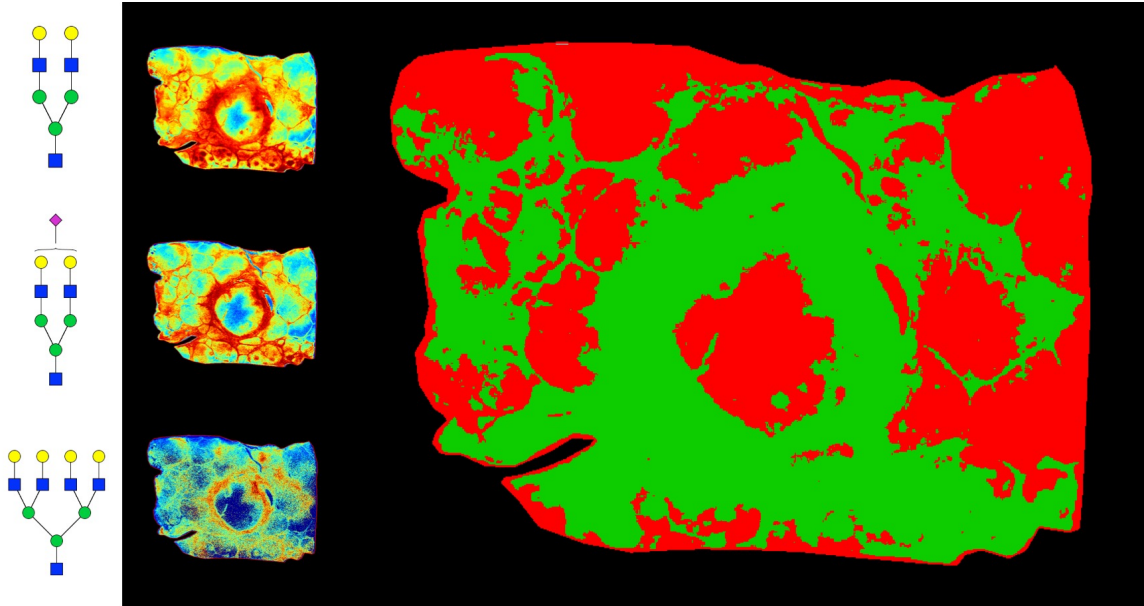
The imaging of the core fucosylated N-glycans was done exactly as described in section 4.3.4.

### **6.2.3. Segmentation Analysis**

After imaging the core fucosylated N-glycans, regions of core fucosylation were established using segmentation analysis. Segmentation was performed using Bruker SCiLS 2017 software (Billerica, MA). Segmentation analysis is a program within SCiLS Lab that allows for parsing specific features found in multiple areas of the tissue and hierarchically clustering them based on these features. Figure 40 demonstrates an example of this segmentation analysis. Briefly, the segmentation analysis was done under the parameters of weak denoising, bisecting k-means, and Manhattan distribution metric. The results were then parsed into two main groups: those containing core fucosylation and those that did not. The results that were achieved by this segmentation are demonstrated in Figure 41. By overlaying this information with the imaging data



**Figure 40: Example Segmentation Analysis.** Example of segmentation analysis dictation regions of similarity as compared to normal (A), cirrhotic (B), and HCC (C) tissues, as well as associated hematoxylin and eosin stains with 1x on the left and 10x magnification on the right.



**Figure 41: Segmentation of Endo F3 Applied HCC Tissue.** Segmentation analysis done following imaging of tissue applied with Endo F3. All components of the analysis were done with core fucosylated associated masses, with red indicating regions where core fucose was not present, while green indicates regions of the tissue in which core fucosylated N-glycans were found.

acquired, regions within the tissue containing core fucosylation were identified and excised from the tissue.

#### **6.2.4. Tissue Extraction and Enrichment**

Once the regions of tissue that contained the core fucosylation were identified via segmentation analysis, the tissue of interest was extracted. This was done by precisely scraping large tissue sections that were not of interest (tumor tissue containing core fucosylated N-glycans) off the slide with a straight edge razor. Next, the regions of interest were circled using a glass scratching pen. The rest of the tissue was then scraped away using the tip of the straight edge razor, leaving only tissue regions of interest left on the slide. Finally, the regions of interest were scraped from the slide and placed in an Eppendorf tube.

#### **6.2.5. Tryptic Digestion**

In-solution tryptic digest was performed as previously described with minor modifications [319]. As the tissue had already undergone antigen retrieval for imaging, this process was skipped in solution. The tissue was brought up in 50 $\mu$ L of 25mM ammonium bicarbonate (AMBIC), ensuring that the pH was between seven and eight. Next, 50 $\mu$ L of trifluoroethanol (TFE) was added and the tissue solution was sonicated in a bath for 15 minutes, ensuring that most of the tissue was dissolved in the solution. The sample was then spun down briefly, parafilm-shut and incubated on a thermoshaker at 60°C and 300rpm.

Next, Dithiothreitol (DTT) and tris-2(-carboxyethyl)-phosphine (TCEP) were added to create an end concentration of 25mM DTT and 10mM TCEP (1 $\mu$ L of 1M TCEP and 5 $\mu$ L of 500mM DTT, stock solutions both in AMBIC). The sample was then parafilmmed shut and incubated at 60°C for 30 minutes. Following the incubation, the sample was cooled to room temperature and iodoacetamide (IAA) was added to make a final IAA concentration of 50mM (approximately 10.6 $\mu$ L of a 500mM stock solution in AMBIC). The sample was then covered with aluminum foil and allowed to incubate in the dark for 40 minutes.

Following the IAA incubation, 50mM AMBIC was added to reduce the total concentration of TFE to 10 percent of the total solution (383.4 $\mu$ L of 50mM AMBIC). This solution was then sonicated for five minutes and briefly spun down. Finally, 2 $\mu$ g of trypsin was added, then the sample was parafilmmed shut and incubated overnight at 37°C. From here, the sample could be stored at -20°C until further proteomic analysis was ready.

#### **6.2.6. Protein Quantification**

To ensure that enough protein was obtained from the isolated tissue sections, a protein quantification assay was performed to determine total protein amount in the sample. Protein quantification was vital, as the sample clean up required a minimum of 10 $\mu$ g of total protein to ensure correct proportions. The sample was quantified using the reducing agent-compatible Pierce Microplate BCA Protein Assay Kit from Thermo Scientific (Catalog #23252). Protocol was

followed as provided by the kit; the results were read on the microplate reader and protein amounts were determined based off the standard curve created. Once the protein concentration was determined to be at least 10µg, the sample then proceeded to clean-up via ZipTip.

### **6.2.7. Sample Clean-up**

The sample was then de-salted via C18 ZipTip cartridges, capable of binding a total of 10µg peptide per sample. Briefly, the sample was first dried in the speed vacuum and brought back up in 30µL of 0.1 percent trifluoroacetic acid (TFA). Next, 50µL of the four necessary solutions were created in separate tubes. First was the wetting solution of 75 percent acetonitrile (ACN) and 0.1 percent TFA, then the equilibrium solution of 0.1 percent TFA, followed by the wash solution of 0.1 percent TFA and finally the elution solution of 75 percent ACN and 0.1 percent TFA.

Once all the solutions were made, the ZipTip was hydrated using 10µL of the wetting solution, solution was then discarded, and repeated two more times. Next, the tip was then equilibrated using 10µL of the equilibrium solution, three times.. Then the peptides were bound to the C18 by pulling up 10µL of the sample and slowly pipetting back down for a total of 30 times, careful to not introduce any air bubbles to allow for maximum binding. After the 30 cycles, the sample solution was then dispensed back into the sample tube.

Next, the tip was washed three times using 10µL of the washing solution. Finally, the sample was eluted from the tip using the elution solution. 10µL of

elution solution was carefully pulled into the tip as slowly as possible and dispensed into a clean tube, again with caution to not introduce any air into the tip, as this would decrease elution quantity. This process was then repeated a minimum of three times to ensure full elution of the peptides. The sample was then dried down in the speed vacuum and stored at -20°C for further proteomic analysis.

### **6.2.8. Orbitrap Fusion Lumos Analysis**

Peptides were separated and analyzed on an EASY nLC 1200 System (ThermoFisher) in-line with the Orbitrap Fusion Lumos Tribrid mass spectrometer (ThermoFisher) with instrument control software version 4.2.28.14. 2µg of tryptic peptides were pressure loaded at 1,180 bar and peptides were separated on a C18 reversed phase column (Acclaim PepMap RSLC, 75µm x 50cm (C18, 2µm, 100 Å) ThermoFisher) using a gradient of 5 percent to 40 percent B in 180 min (Solvent A: 5 percent acetonitrile/0.1 percent formic acid; Solvent B: 80 percent acetonitrile/0.1 percent formic acid) at a flow rate of 300nL/min with a column heater set to 50°C.

Mass spectra were acquired in data-dependent mode with a high resolution (60,000) FTMS survey scan, mass range of 375-1575 m/z, followed by tandem mass spectra (MS/MS) of the most intense precursors with a cycle time of 3 seconds. The automatic gain control target value was 4.0e5 for the survey MS scan. Fragmentation was performed with a precursor isolation window of 1.6 m/z, a maximum injection time of 22 ms, and HCD collision energy of 35 percent.

Monoisotopic-precursor selection was set to “peptide”. Apex detection was not enabled. Precursors were dynamically excluded from resequencing for 30 seconds and a mass tolerance of 10 ppm. Precursor ions with charge states that were undetermined or greater than 5 were excluded.

### **6.2.9. Protein and Peptide Identification**

Proteome Discoverer version 1.4.0.288 (ThermoFisher) was used to determine protein and peptide identification from the sample run. Briefly, the raw data underwent two major processing nodes: Spectrum Selector and Sequest HT. For the Spectrum Selector, lower limits were 15 minutes retention time and precursor mass of 350 Da and upper limits were 200 minutes with a precursor mass of 5000 Da. Signal to noise (S/N) ratio for the fourier-transform was 1.5, with the MS2 activation set at higher-energy collision-dissociation. For the Sequest HT node, the protein database searched was SPTR\_092718\_Human, with the enzyme being Trypsin, a max missed cleavage site of 2, minimum peptide length of 7 and maximum peptide length of 144. The mass tolerances were set at 20 ppm for the precursor and 0.02 Da for the fragment. Finally, three dynamic modifications and one static modification was included in the search. For the dynamic modifications, oxidation (+15.995 Da M), deamidated (+0.984, N or Q), and HexNAc+dHex (+349.137 Da, N) were examined, while the static modification was carbamidomethyl (+57.021 Da, C). This node searched a total of 173,778 sequences. This search resulted in a total of 117,992 search inputs with 1146 protein groups, 8992 merged proteins, and 3826 peptides.

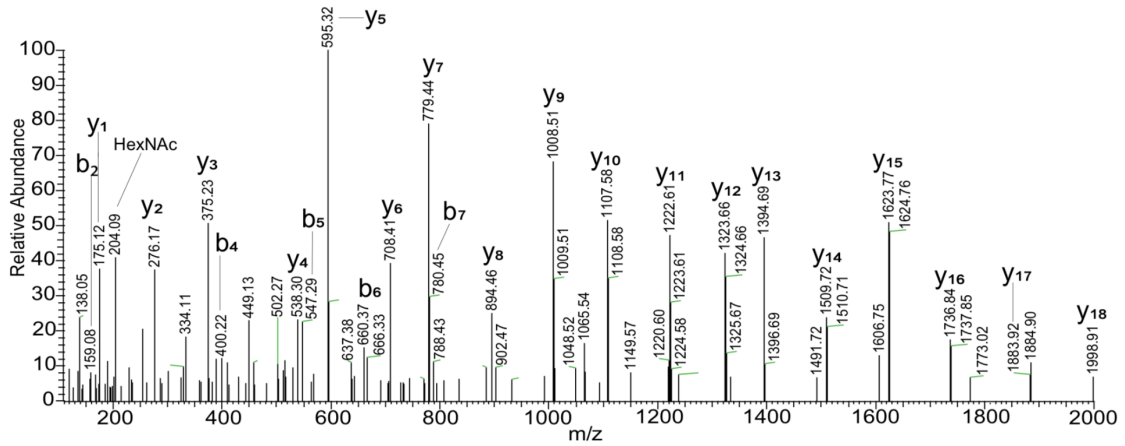


### 6.3. Results

Initially, following the analysis, representative spectra containing the modification of interest were pulled and annotated according to their b and y ions left from the HCD fragmentation pattern. Figure 42 demonstrates an example of one peptide. This specific peptide contains 37 amino acids in the sequence ASLQFLQNYTALASAVDAMDFINDATDVNDALGYVTR (corresponding to Collagen alpha-1(VI)) with the Endo F3 modification occurring on the asparagine in the eighth position of the sequence. The top of the figure represents the sequence as well as the corresponding b and y ion fragments resulting from the fragmentation of this peptide sequence. The bottom part of the figure demonstrates the relative abundance of these ions in the spectra, sorted by mass to charge ratio. For example, the mass value for  $y_1$  is 175.12 m/z, corresponding to the amino acid of arginine, while the mass value for  $y_2$  is 276.17 m/z, corresponding to both arginine and threonine, and so on. These spectra are what allows for the assignment of amino acids per retention peak via LC/MS, allowing for proper peptide identification while still maintaining the modification.

Following the assignment of peptide amino acids, the data was then processed to return a protein identification search, where peptides containing the tag modification left behind by Endo F3 cleavage were assigned to proteins, and functional protein groups were established. Table 10 provides a list of all peptides found to contain the core fucose modification (35 total tagged peptides), as well as the peptide spectrum matches, number of proteins associated with the peptide, the protein groups, the accessions, where the modification occurs as

b 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24  
 A - S - L - Q - F - L - Q - N(HexNAcFuc)- Y - T - A - L - A - S - A - V - D - A - M(ox)- D - F - I - N - D  
 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14  
  
 25 26 27 28 29 30 31 32 33 34 35 36 37  
 A - T - D - V - N - D - A - L - G - Y - V - T - R  
 13 12 11 10 9 8 7 6 5 4 3 2 1 y



**Figure 42: Annotation of Representative MS/MS Spectrum.**

Table 10: Protein Identifications Containing Core Fucose Modification.

A2	Sequence	# PSMs	# Proteins	# Protein Groups	Protein Group Accessions	Modifications	ΔCn	Xcorr Charge	MH+ [Da]	ΔM [ppm]	m/z [Da]	RT [min]	# Missed Cleavages
High	ASIQQTNTALASAVDAmDFINDADVDNALGVYTR	28	3	1	A0A087X055	N8(HexNAc1)Hex(1); M19(Oxidation)	0.0000	4.18	3	4359.06552	187.10	0	
High	IHYLYLQINFTLPEVFSQATGLR	22	1	1	P51888	N21(HexNAc1)Hex(1)	0.0000	3.76	3	3429.72800	143.94	0	
High	QLIALQINFTLPEVFSQATGLR	2	7	1	P12111-2	N4(HexNAc1)Hex(1)	0.0000	3.43	3	2732.50412	109.14	0	
High	LGLSFSASVANGSLANTPLR	18	5	1	P07585-3	N6(HexNAc1)Hex(1)	0.0000	3.36	3	2732.37925	106.52	0	
High	IHYLYLQINFTLPEVFSQATGLR	24	1	1	P51888	N8(Deamidated); N21(HexNAc1)Hex(1)	0.0000	3.33	3	3430.72226	114.24	0	
High	GONQRFHFLVQAVSFGACVAVSDLRK	2	6	1	BADNDA	N12(HexNAc1)Hex(1); C23(Carbamidomethyl)	0.0000	3.26	4	3958.98818	148.42	1	
High	LGLSFSASVANGSLANTPLR	10	5	1	P07585-3	N6(HexNAc1)Hex(1); N13(Deamidated)	0.0000	3.25	3	2733.36820	140.38	0	
High	LGLNKRHDIENSLANPR	4	2	1	Q9BANN1	N6(HexNAc1)Hex(1)	0.0000	3.14	3	2558.33731	109.62	0	
High	QAPGGHLSLGFVWELGLAPAGADLNLESDDVQLNLR	1	1	0	Q00206-3	C10(Carbamidomethyl); N39(HexNAc1)Hex(1)	0.0977	3.14	5	4712.23493	65.89	1	
High	YDLSPNGVTWISSFLGLEQLHDFHSHLK	2	4	1	Q9Y490	N15(Deamidated); N31(HexNAc1)Hex(1)	0.0000	3.06	5	4160.01203	63.60	0	
High	VEHSVALPAmrSASGFERFQVSSPPAQQTSGQMRK	2	2	1	Q9Y490	M12(Oxidation); N21(HexNAc1)Hex(1); O31(Deamidated)	0.0000	3.05	5	4654.20515	63.60	0	
High	VGNQAVVTVGRVPPYSLQLTWSENGVQQR	2	2	0	Q9Y490	N6(HexNAc1)Hex(1); C9(Carbamidomethyl); Q10(Deamidated); C30(Carbamidomethyl)	0.0163	3.01	5	4160.01203	72.48	2	
High	GLKRPALTPASITFTGLEAPR	5	15	1	A0A14078Y3	N6(HexNAc1)Hex(1)	0.0000	3.00	3	2705.40451	100.16	1	
High	TALQPGQQQSDSNNKPKGSDSSGR	1	1	0	B3K5AG	O5(Deamidated); N20(HexNAc1)Hex(1)	0.0810	2.95	5	3420.62025	59.6	1	
High	nLQAQEVETQNLKPGSDSNDHLQKYLKLELTSR	1	1	1	P02458-1	N1(HexNAc1)Hex(1); C25(Carbamidomethyl)	0.0000	2.94	5	4653.16024	64.78	2	
High	GAPKPRGRDGEPTGPGPGPGPGPGPLGWFNAAMGGFERK	1	2	1	B4DZ8-1	N17(HexNAc1)Hex(1)	0.0000	2.91	5	4654.22071	79.17	2	
High	TAKESGGYEPFLTWFMQWVFLTATLPHVTLK	1	1	1	C9IPW9	C30(Carbamidomethyl); N33(HexNAc1)Hex(1); Q34(Deamidated)	0.0000	2.88	5	4027.05549	63.78	1	
High	EITLSRQTLNPTFTTRFNPWPLRSVLSYK	1	3	1	C9IPW9	N22(HexNAc1)Hex(1)	0.0000	2.81	5	4027.05549	5.99	1	
High	qNGIPRSTAFHGTSSNGALTKADQLPmVMSVR	3	4	1	Q9UBZ9-3	Q1(Deamidated); M15(HexNAc1)Hex(1); O28(Deamidated); C30(Carbamidomethyl); N36(Deamidated)	0.0000	2.80	5	4668.24293	2.63	1	
High	KHWNSALQMPDPLVTLGLASDQLDmVMSVR	1	1	0	Q7LEZ2	M26(Oxidation); N29(HexNAc1)Hex(1)	0.0414	2.78	5	4071.94624	7.12	2	
High	VDAASLQVWAKRGLGIPLSLSEVGR	2	1	0	Q7LEZ2	O9(Deamidated); N17(HexNAc1)Hex(1)	0.0111	2.67	4	4398.07583	7.32	1	
High	DVLMQDFVSEFFSFGSALTAHFFDFRK	1	6	1	Q7LEZ2	N21(HexNAc1)Hex(1); O29(Deamidated)	0.0000	2.64	4	4398.07583	3.85	1	
High	HSWYHGPVSRRAEHLSSNGSFLVRESSEPGQR	1	1	0	Q7LEZ2	N11(HexNAc1)Hex(1)	0.0000	2.64	4	4398.07583	8.57	1	
Medium	mIENSLSLPTFR	6	11	1	A6N1G9	Q10(Deamidated); M15(Oxidation); N25(HexNAc1)Hex(1)	0.0152	2.59	4	3326.67885	8.40	2	
High	ASIQQTNTALASAVDAmDFINDADVDNALGVYTR	4	3	1	A0A087X055	M1(Oxidation); N4(HexNAc1)Hex(1)	0.0000	2.55	3	1942.97257	2.42	0	
High	FLHSGAGAGVLTSGDAGMVAWR	1	1	1	Q01813	Q7(Deamidated); N8(HexNAc1)Hex(1); M19(Oxidation)	0.0000	2.52	4	3176.57436	4.73	1	
Medium	GENGVGTPGVGAAGPAGPGPGPAGSRGSGGPGMTEFFGAAGR	1	2	1	A0A087MTA8	N25(HexNAc1)Hex(1)	0.0000	2.51	4	4383.07290	0.72	1	
High	VLPVAVTDVGLVR	19	2	1	P35555	N5(HexNAc1)Hex(1); C10(Carbamidomethyl)	0.0000	2.49	2	2025.07264	3.30	0	
High	AKLDFEALFEPTPESSSR	2	2	1	B2R9V7	N17(HexNAc1)Hex(1)	0.0000	2.48	3	2633.28330	8.15	1	
High	RWGTPAONTGTEPVSVEWIK	1	4	1	Q8NDV7-2	N13(HexNAc1)Hex(1); N20(Deamidated)	0.0000	2.48	3	2663.27140	4.85	1	
High	IHYLYLQINFTLPEVFSQATGLR	5	1	1	P51888	N8(Deamidated); N9(Deamidated); N21(HexNAc1)Hex(1)	0.0000	2.48	3	3431.70810	5.45	0	
High	mIITLSQVLEAK	6	4	1	P12110	N1(HexNAc1)Hex(1); M2(Oxidation)	0.0000	2.40	2	1732.82781	4.62	0	
Medium	nTSAASTAQVATELRR	1	3	1	P52268	N1(HexNAc1)Hex(1)	0.0000	2.27	3	2396.18131	1.04	1	
High	ATPEAAHSELAALR	2	2	1	K7EKF6	N7(HexNAc1)Hex(1)	0.0000	2.16	2	1833.91033	1.43	0	

well as any other modifications, the cross-correlation score, overall mass, charge, mass to charge ratio, retention time, and mixed cleavages. All peptide sequences had a cross-correlation value above 2 and most had a relatively low normalized score difference (all but one were less than 0.05), demonstrating that while the sample was not enriched for the core fucosylated glycopeptides any further than the segmentation analysis, confidence in these modified peptides was high.

Finally, the accession numbers associated with the searched peptides were used to determine likely protein identification. Table 11 demonstrates the accession number associated with a modified peptide, the protein corresponding to that accession number, as well as the molecular function of the protein. From the table, it is clear that a vast majority of the peptides found were associated with collagen proteins, specifically collagen alpha-1(VI), while many of the other glycoproteins found were connected with other types of glycosylation, such as glycosaminoglycans found on Biglycan, Decorin, or Asporin, as well as other extracellular matrix-associated proteins, such as Prolargin. These findings are consistent with what was expected regarding this specific tissue, as many of the regions of core fucosylation were along these “collagen highways,” implying that many of these core fucosylated proteins should indeed be associated with collagen binding or the extracellular matrix.

Interestingly, while many glycoproteins were as expected, there were also some protein groups that were surprising, such as redox-based regulation, DNA binding, and even synaptic regulation. Further studies will need to be done to

**Table 11: Primary Biological Component or Pathway Containing Core Fucose Modification.**

ACCESSION NUMBER	PROTEIN	MOLECULAR FUNCTION
A0A2P9AJL7	Acetolactate Synthase Large Subunit ILVX	Acetolactate Synthase Activity
K7EKF6	Angiopoietin-related protein 6	Angiogenesis Induction / Cell Proliferation
Q9BXN1	Asporin	Calcium Ion/Collagen Binding
Q01813	ATP-dependent 6-phosphofructokinase, platelet type	AMP/ATP Binding / Cadherin Binding
A6NLG9	Biglycan	Collagen Fiber Assembly
B4DND4	cDNA FLJ50588 (Gamma glutamyltransferase 5)	Glutathione Hydrolase Activity / Transferase
P12110	Collagen alpha-1(VI)	Collagen Binding
A0A087X0S5	Collagen alpha-1(VI)	Collagen Binding
A0A087X0S5	Collagen alpha-1(VI)	Collagen Binding
A0A087X0S5	Collagen alpha-1(VI)	Collagen Binding
A0A087VMTA8	Collagen alpha-2(I)	ECM Constituent
P12111-2	Collagen alpha-3(VI)	ECM Constituent / Serine-Type Endopeptidase Inhibitor
P02458-1	Collagen alpha-1(II)	ECM Constituent / Proteoglycan Binding
P07585-3	Decorin	Collagen Binding / ECM Binding / GAG Binding
P07585-3	Decorin	Collagen Binding / ECM Binding / GAG Binding
Q9UBZ9-3	DNA repair protein REV1	Collagen Binding / ECM Binding / GAG Binding
A8K979-2	ERI1 exoribonuclease 2	Collagen Binding / ECM Binding / GAG Binding
P35555	Fibrillin-1	Damaged DNA Binding
Q96PQ7-4	Kelch-like protein 5	3'-5'-Exoribonuclease Activity / Nucleic Acid Binding
Q7LE22	KIAA0589 (Fragment)	ECM Constituent / Heparin + Integrin Binding
P55268	Laminin subunit beta-2	Actin Binding
C9JPM9	Leucine-rich repeat transmembrane neuronal protein 1 (fragment)	ATP Binding
A0A087WUJ78	Nance-Horan Syndrome Protein	ECM Constituent / Integrin Binding
P51888	Prolargin	ECM Organization / Negative + Positive Regulation
P51888	Prolargin	Cell Differentiation
P51888	Prolargin	ECM Constituent / Heparin Binding
B2R9V7	Superoxide dismutase	ECM Constituent / Heparin Binding
Q9Y490	Talin-1	ECM Constituent / Heparin Binding
A0A140T8Y3	Tenascin-X	Metal Ion Binding
O00206-3	Toll-Like Receptor 4	Actin/Cadherin/Integrin/Vinculin Binding
Q5T4D3-4	Transmembrane and TPR repeat containing protein 4	Collagen Fibril Organization
Q8NDV7-2	Trinucleotide repeat-containing gene 6A protein	Signaling Receptor Activity
P18206-2	Vinculin	Mannosyltransferase Activity
		RNA Binding
		Actin/Catenin/Cadherin Binding

determine the accuracy or reliability of these glycoproteins' association with HCC, though the initial results are promising in demonstrating new and more specific glycoprotein changes associated with HCC progression and early diagnosis.

#### **6.4. Discussion**

The work above demonstrates a highly promising start for enhanced biomarker identification for liver cancer specifically. Traditional serum glycoproteomic analysis has been examined before in HCC, as well as the use of endoglycosidases to determine core fucosylation in hepatic serum via the HexNAc+Fuc tag [144, 152, 320-323], however this work effectively demonstrates the first use of endglycosidase F3 activity via an imaging platform for enhancing glycoproteomic targets within a sample without further enrichment. While further enrichment would undoubtedly yield a larger result of more specific proteins due to the exclusion of less abundant peaks in an LC-MS/MS experiment, the workflow above demonstrated a promising start.

In terms of findings within this chapter, ultimately the glycoproteins found were novel in terms of enhancement for tissue-specific glycoproteins; however, further studies would need to be done to compare these findings with those in serum. For example, while the initial experiments yielded promising results, many known and well established glycoproteins found in HCC serum, such as alpha-fetoprotein, were absent from this list, implying that this pool of glycoproteins would be separate from those found in circulating serum [324]. While this is less

ideal for the identification of a more specific biomarker, these findings could still open a pathway for other drug or therapeutic targets within HCC that were previously unknown from lack of direct HCC tissue analysis. Interestingly, the upregulation of ECM-related glycoproteins is consistent with newly published studies, such as one that found that patients with HCC had significantly higher collagen-III levels and a marked increase in the collagen-III/MMP-1 ratio (CMR) [325]. The continued study of these fucosylated ECM glycoproteins could provide a functional link between HCC progression and earlier diagnosis, especially when examined in a panel with other known markers such as serum AFP levels.

However, before absolute certainty could be assigned to glycoproteins found via this method, the strategy would require further refinement. As this was only one tissue sample, pools would be relevant to examine through this method to accurately assess whether these glycoproteins are patient-specific or generally found throughout the disease state. Cirrhotic controls would also be essential to assess presence of these glycoproteins as compared to a “normal” control state for patients.

Future directions with this work, beside further optimization for the methodology, would be application to other disease states. While HCC is a logical starting choice, as serum is hepatic in origin, many other diseased tissue types have not been examined in such a specific manner, despite core fucosylation being implicated in a multitude of disease states. With application of this method to other disease states, new glycoprotein targets could be implicated

in disease progression, initiation, or metastasis, opening a new field of potential therapeutics for a variety of disease types.



# **Chapter 7: Conclusions, Limitations, and Future Studies**

## **7.1. Overall Findings**

With HCC research remaining primarily in the realm of serum analysis and showing great promise in the field of glycomics, the need for more extensive examination of HCC tissue is greatly apparent. The work presented in this dissertation is outlined by three key areas: 1) examining the changes in N-glycosylation of HCC tissue specifically, determining an upregulation in fucosylation and complex glycosylation; 2) developing methods to further characterize N-glycan isoforms while maintaining imaging applications, specifically for determining core versus outer arm fucosylation and differentiation of sialic acid linkages; and 3) applying these characterizations and the resulting tags and/or mass shifts combined with the imaging methodology for enhanced glycoproteomic analysis of core fucosylated glycoproteins. Below, the broader implications, conclusions, limitations, and future directions are discussed for each key area.

## **7.2. Changes in N-Glycans of Hepatocellular Carcinoma via MALDI-IMS**

### **7.2.1. Conclusions**

At the completion of Specific Aim 1, the changes in N-glycan distribution and abundance were examined within HCC tissue specifically, as well as compared to normal and cirrhotic tissue. The conclusions of chapter three are as follows: 1) there were more than 60 N-glycans upregulated in HCC tissue as opposed to normal or cirrhotic (over the 138 patient samples examined), 2) there

were two major classes of upregulated glycans: increased fucosylation and increased branching, 3) fewer sialylated species of N-glycans were found than expected, and 4) increased branching and fucosylation are inversely associated with patient survival.

The first conclusion demonstrates an overall examination of the N-glycome for direct HCC tissue. While tissue analysis of HCC tissue via MALDI-IMS had been done previously in the lab [195], this research was novel for examining HCC tissue in a larger cohort, allowing for further conclusions about disease-state trends rather than purely associating results with a single patient-specific tissue. The two separate tissue microarrays, along with the larger tissue sections previously examined, gave a much more comprehensive understanding of cancerous, cirrhotic, and normal tissue types, providing increased accuracy in determining N-glycosylation trends.

The findings from these patient imaging experiments revealed two major types of N-glycosylation within HCC tissue. The first, fucosylation, was overexpressed in 96 percent of patients, with a range of 1 to 33 fucosylated N-glycans over-expressed in each patient. This finding was consistent with the literature, showing that fucosylation is related to HCC progression and was validated in a large sample cohort. The second type of N-glycan modification was that of increased branching, where tetra-antennary glycan structures were found primarily in HCC tissue alone. This is also consistent with previous findings, as the enzyme responsible for tetra-antennary N-glycans (MGAT5) has been

associated with many cancers previously through both hTERT and the Ras/Raf pathway [326]

Similarly, the associated patient survival data was key in determining the integral findings of this chapter. With these purchased tissue microarrays, we were also provided patient survival data, as well as grading, staging, and etiology of each patient sample. This, in turn, allowed for a larger variety of analyses, providing an in-depth look at the role N-glycans play in HCC progression and survival. While grading and staging showed no correlation with N-glycan expression, an increase in N-glycan branching and fucosylation was shown to have an inverse effect on patient survival; the highly-expressed presence of these N-glycan types reduced overall survival time by an average of 19 months. This novel finding is one of the first key links between N-glycan expression and overall survival for patients with HCC.

### **7.2.2. Limitations and Future Research**

As discussed in chapter three, the largest limitation with these imaging experiments was our inability to determine direct linkages within the N-glycans, providing doubt in determining what N-glycan structures exactly were associated with the HCC tissue. For example, while fucosylation was demonstrated in 96 percent of patients, we were unable to determine whether these N-glycans were core fucosylated or outer arm fucosylated, as either could have varying implications within the disease state. This limitation was in part addressed throughout chapter four, although further work would need to be done to ensure

validation of this method (more information to follow in section 7.2.2.) Similarly, the lack of sialylation found within the tissue was problematic, as the resulting N-glycome did not accurately represent the true N-glycome of HCC. As discussed in chapter five, this could result in a misrepresentation of N-glycan distribution, although ultimately the amount of sialylation would not affect the general trends of increased fucosylation and branching found in chapter three.

One other limitation found when examining the larger cohort of patient data was the sheer volume of glycan heterogeneity within each patient. While we were clearly able to establish trends in the glycosylation patterns, the individual glycan profile still widely varied between patients. This made it difficult to accurately assess specific N-glycan changes, relying on broad trends instead. Other members of the Mehta lab are working to address this, but briefly, this could be combatted through examination of the N-glycome for different liver genetic subtypes. If each subtype is analyzed for changes in glycosylation as compared to any normal tissue of the same genetic subtype, more conclusions for specific changes of glycosylation can be attributed to different subtypes. This could ultimately lead to a more efficient method for classification of each genetic subtype of patient, as well as increase therapeutic effects. By knowing an HCC patient's genetic subtype, different therapies could be used to combat the type of progression or disease infiltration (as discussed in chapter one).

Finally, one future direction that could improve this research for future biomarker studies is to compare these tissue findings to matching serum samples. While many studies have been done to examine serum glycosylation

trends for HCC as compared to cirrhotic or normal, the literature is lacking direct tissue analysis corresponding to associated patient serum. If matched tissue and serum samples could be obtained, a side-by-side analysis could reveal more direct trends associated with HCC, as well as match trends in serum glycosylation to direct HCC tissue glycosylation changes. This could also assist in determining more specific trends for the genetic subtypes of HCC, providing more specific analysis of the genetic disease state from a readily obtainable and less invasive fluid biomarker such as serum.

### **7.3. Enzymatic and Chemical Characterization of N-Glycans for MALDI-IMS**

#### **7.3.1. Conclusions**

Looking at the enzymatic and chemical derivatization methods established in chapters four and five, several major conclusions can be drawn regarding the efficiencies of these methods, as well as their applicability to HCC and relevance in a biological context. For the enzymatic characterization using Endo F3, the primary conclusions are as follows: 1) we were able to accurately apply Endo F3 as an enzymatic digestion on tissue to preferentially cleave core fucosylated N-glycans, and 2) we were able to establish this methodology in a sequential and combined fashion to allow for more in-depth analysis of isomeric linkages of fucosylation for N-linked glycans.

The first conclusion regarding application of Endo F3 to tissue for identifying core fucosylated N-glycans is completely novel; it is the first application of this enzyme to tissue sections for imaging mass spectrometry to

examine differential cleavage. With the enzyme cleaving at a different position on the N-glycan, the mass shift enables visualization of these core fucose structures on tissue while still maintaining all benefits afforded through imaging.

Additionally, preference to core fucosylation was established via HPLC analysis and throughout the method development process, allowing this new method for tissue imaging to be widely applied throughout our lab and others.

Along with the application of the singular enzyme to tissue for analysis, the second conclusion effectively demonstrated that Endo F3 can be used both sequentially and concurrently with PNGase F, resulting in less sample preparation and usage, while still maintaining the increased evaluation of the fucosylation linkages on N-glycans. Sequentially, it was shown that we can apply Endo F3 first, analyze via MALDI imaging, then wash it away and apply PNGase F to the same tissue, enabling a direct comparison between core fucosylation to outer arm fucosylation on the same sample. Serendipitously, it was also discovered that using both Endo F3 and PNGase F concurrently, provided the Endo F3 concentration is relatively low compared to the PNGase F concentration, works exceedingly well in examining both sets of data within the same imaging experiment. This is even more beneficial than the former as this removes any variability from experiment to experiment, allowing for more direct quantitative analyses between the core and outer arm fucosylation.

As for the chemical derivatization of sialic acids on tissue for imaging analysis, two major conclusions were drawn regarding this established methodology on HCC tissues: 1) it was established that the amidation-amidation

reaction is overall more effective at stabilization, with an increased prevalence of multi-sialylated species and 2) while  $\alpha$ 2,6-linked sialic acids were overall more abundant in the tumor tissue of HCC, particularly with fucosylation involved, there was no direct evidence supporting specific sialylated species being directly related to tumor versus alternative disease or normal tissue state.

Experiments demonstrated a clear benefit to using amidation over ethyl esterification, with Figure 35 and 37 clearly showing the efficiencies of both these methods on prostate tissue. For non-derivatized versus derivatized in both methods, it is apparent that amidation resulted in a better stabilization for tissue. This is also applicable to the HCC tissue. While there is no data for ethyl esterification on liver tissue specifically, amidation of the liver tissue worked extremely efficiently, allowing us to see the differential mass shift associated with the linkage of the sialic acid.

Interestingly, for the second conclusion, while consistent with findings in the literature, there was no apparent association between sialylation and HCC progression or diagnosis. Figure 39 demonstrated an overall bias toward the  $\alpha$ 2,6 conformation, however when compared to different tissue morphologies provided in the tissue microarray, the overall findings were inconclusive at establishing a link between sialylation and disease progression.

### **7.3.2. Limitations and Future Research**

In terms of the characterization of core fucosylation, one major limitation with the research is the activity of the enzyme. While Endo F3 has been shown to



have a much higher prevalence to core fucosylated structures under the right enzyme conditions, it is also still able to cleave tri-antennary N-glycans without a fucose. This complicates the analysis of core fucosylation, with this methodology not being able to identify that cleaved N-glycans were core fucosylated with an absolute certainty. The confirmation via HPLC and the optimization of the method conditions greatly support the enzymes' cleavage of core fucose, however this fact can still not be ignored.

For future studies regarding the methodology of Endo F3 on tissue, it would be helpful to further validate core fucose preference via lectin histochemistry. With previous work done in the Mehta lab, a mutated recombinant lectin was created that preferentially binds to core fucose, providing an additional confirmation of core fucose [275]. While this lectin is also not absolutely core fucose specific, this could provide further validation when stained on tissue that has been treated with Endo F3. This validation could also occur through a proteomic lens. As discussed in chapter six, the Endo F3 tag left behind is efficient in identifying glycopeptides containing core fucose; however, if this idea was applied to examine just HexNAc tags without the fucose, the efficiency of Endo F3 cleaving tri-antennary glycans without a core fucose could be further examined. This could ultimately lead to a greater understanding of the enzyme activity and cleavage of N-glycans.

Regarding the derivatization of the sialic acids, I believe that further studies are required before any concrete ideas connecting sialylation and liver cancer can be determined. Similar to what was done in chapter three, validation

of sialylation on a patient matched TMA would not only provide more information for a broader patient sense, but also allow for more analysis of sialylation comparing to survival data. As stated above, the overall levels of sialylation were lower than expected, so if the same TMA was examined using these stabilization techniques, the resulting N-glycome could shift, revealing new trends as it pertains to an overall population with HCC, as well as associated survival times, grading, and staging of the tumors.

## **7.4. Enhanced Glycoproteomic Analysis of Hepatocellular Carcinoma Tissues**

### **7.4.1. Conclusions**

From the initial glycoproteomic experiments, the primary conclusions found were 1) imaging with Endo F3 and using this analysis for preferential selection of tissue adds benefit in examining core fucosylated glycoproteins, and 2) the remaining HexNAc+Fuc tag following Endo F3 digestion can effectively be used with tissue for enhanced proteomic analysis of core fucosylated glycoproteins.

When Endo F3 was imaged on tissue, segmentation analysis allowed for differentiating core fucose-containing regions of the tissue from those without major areas of core fucosylation. This, combined with isolation of the regions of interest via manual microdissection, allowed for enhancement of target glycopeptides without any further enrichment. The resulting glycoprotein analysis revealed 47 total peptides with the associated tag of HexNAc+Fuc, leading to

confident identification of 28 glycoproteins associated with core fucosylation. While this does not encompass the full spectrum of glycoproteins with core fucosylation, this is a promising start for early work without further enrichment methods.

#### **7.4.2. Limitations and Future Research**

The major limitation associated with this section of my research is the overall sample size and population and the lack of enrichment within the sample. While promising data was obtained without the need for further enrichment, this would ultimately aid in the most accurate and comprehensive look at the core fucosylated glycopeptides in HCC tumor tissue specifically.

For example, enrichment via lectin, such as Concanavalin A, would enrich for all glycopeptides in solution, thus reducing the background noise of this method and increase identification of less abundant glycopeptides that may have too little intensity to be picked up. Additionally, with this method shown to have promising results, adapting the specific parameters could yield better results. With the Orbitrap Fusion Lumos having capabilities for higher-energy collision dissociation product-dependent electron transfer dissociation, the resulting tag could be utilized more specifically in a product-dependent fashion, thus enhancing the sample for the glycopeptides containing this tag.

In future studies, a more in-depth look at various tissues from multiple sources in one pool would be extremely beneficial in determining the glycoprotein most related to HCC tissue specifically. A study that contained pooled HCC

tissues, cirrhotic tissues, and normal tissues would greatly enhance the results of this experiment in examining the trends of a wider sample population rather than in one individual.

Additionally, it would also be relevant to examine associated serum along with tissue via this method. While others have examined core fucosylated glycoproteins in serum, if patient-matched serum and tissue were examined concurrently, specific proteins in tissue could ultimately be targeted in serum and validated as a potential biomarker candidate in the future. By combining these two orthogonal methodologies, these findings could be applied more efficiently to the field of biomarker discovery and enhancing possible therapeutic targets, as well as elucidate previously unknown underlying mechanisms of HCC.

## **7.5. Conclusions and Final Thoughts**

Overall, the work presented in this dissertation demonstrates a novel examination of N-glycan changes associated within hepatocellular carcinoma and presents novel methodologies for examining isomeric linkages of N-glycans while still maintaining an imaging mass spectrometry analysis platform. Additionally, initial enhancement of targeted glycoproteomics was also presented, showing an improvement from previous literature examples through the combination of both imaging studies and glycoproteomic analysis. Ultimately, this work provides new imaging analyses and a sturdy foundation for identification of glycoproteins associated directly with HCC tumors for increased biomarker possibilities and therapeutic capabilities.

## References:

1. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012*. Int J Cancer, 2015. **136**(5): p. E359-86.
2. Tsuchiya, N., et al., *Biomarkers for the early diagnosis of hepatocellular carcinoma*. World J Gastroenterol, 2015. **21**(37): p. 10573-83.
3. Ryerson, A.B., et al., *Annual Report to the Nation on the Status of Cancer, 1975-2012, featuring the increasing incidence of liver cancer*. Cancer, 2016. **122**(9): p. 1312-37.
4. Mehta, A., H. Herrera, and T. Block, *Glycosylation and liver cancer*. Adv Cancer Res, 2015. **126**: p. 257-79.
5. America, C.T.C.o. *Liver Cancer Types*. 2020 [cited 2020 April 2].
6. Keng, V.W., D.A. Largaespada, and A. Villanueva, *Why men are at higher risk for hepatocellular carcinoma?* Journal of hepatology, 2012. **57**(2): p. 453-454.
7. Bolondi, L., *State of the art: hepatocellular carcinoma*. Future Oncology, 2014. **10**(15s): p. 1-6.
8. Zhu, R.X., et al., *Epidemiology of Hepatocellular Carcinoma in the Asia-Pacific Region*. Gut and liver, 2016. **10**(3): p. 332-339.
9. Ryerson, A.B., et al., *Annual Report to the Nation on the Status of Cancer, 1975-2012, featuring the increasing incidence of liver cancer*. Cancer, 2016. **122**(9): p. 1312-1337.
10. Pang, T.C. and V.W. Lam, *Surgical management of hepatocellular carcinoma*. World J Hepatol, 2015. **7**(2): p. 245-52.
11. Heimbach, J., et al., *Aasld guidelines for the treatment of hepatocellular carcinoma*. Hepatology, 2017.
12. Mehta, A., H. Herrera, and T. Block, *Glycosylation and liver cancer*. Adv Cancer Res, 2015. **126**: p. 257-279.
13. Hoshida, Y., et al., *Integrative Transcriptome Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma*. Cancer Research, 2009. **69**(18): p. 7385.
14. Lee, S.H., et al., *Molecular Subtypes and Genomic Signatures of Hepatocellular Carcinoma for Prognostication and Therapeutic Decision-Making*, in *Hepatocellular Carcinoma: Translational Precision Medicine Approaches*, Y. Hoshida, Editor. 2019, Springer International Publishing: Cham. p. 109-123.
15. Cancer Genome Atlas Research Network. Electronic address, w.b.e. and N. Cancer Genome Atlas Research, *Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma*. Cell, 2017. **169**(7): p. 1327-1341.e23.
16. Block, T.M., et al., *Molecular viral oncology of hepatocellular carcinoma*. Oncogene, 2003. **22**(33): p. 5093-107.
17. Sun, V.C.-Y. and L. Sarna, *Symptom management in hepatocellular carcinoma*. Clinical journal of oncology nursing, 2008. **12**(5): p. 759-766.

18. Villanueva, A., et al., *Genomics and signaling pathways in hepatocellular carcinoma*. Semin Liver Dis, 2007. **27**(1): p. 55-76.
19. Llovet, J.M., M. Schwartz, and V. Mazzaferro, *Resection and liver transplantation for hepatocellular carcinoma*. Semin Liver Dis, 2005. **25**(2): p. 181-200.
20. Petrick, J.L., et al., *International trends in liver cancer incidence, overall and by histologic subtype, 1978-2007*. Int J Cancer, 2016. **139**(7): p. 1534-45.
21. El-Serag, H.B., *Hepatocellular Carcinoma*. New England Journal of Medicine, 2011. **365**(12): p. 1118-1127.
22. Chen, C.-J., et al., *Hepatitis B virus DNA levels and outcomes in chronic hepatitis B*. Hepatology, 2009. **49**(S5): p. S72-S84.
23. Donato, F., et al., *Alcohol and Hepatocellular Carcinoma: The Effect of Lifetime Intake and Hepatitis Virus Infections in Men and Women*. American Journal of Epidemiology, 2002. **155**(4): p. 323-331.
24. Lok, A.S., et al., *Incidence of Hepatocellular Carcinoma and Associated Risk Factors in Hepatitis C-Related Advanced Liver Disease*. Gastroenterology, 2009. **136**(1): p. 138-148.
25. Hartke, J., M. Johnson, and M. Ghabril, *The diagnosis and treatment of hepatocellular carcinoma*. Semin Diagn Pathol, 2017. **34**(2): p. 153-159.
26. Byrne, C.D. and G. Targher, *NAFLD: A multisystem disease*. Journal of Hepatology, 2015. **62**(1, Supplement): p. S47-S64.
27. Wu, H.C. and R. Santella, *The role of aflatoxins in hepatocellular carcinoma*. Hepatitis monthly, 2012. **12**(10 HCC): p. e7238-e7238.
28. Resson, H.W., et al., *Multi-omic approaches for characterization of hepatocellular carcinoma*. Conf Proc IEEE Eng Med Biol Soc, 2016. **2016**: p. 3437-3440.
29. Whittaker, S., R. Marais, and A.X. Zhu, *The role of signaling pathways in the development and treatment of hepatocellular carcinoma*. Oncogene, 2010. **29**(36): p. 4989-5005.
30. Hoshida, Y., et al., *Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma*. Cancer Res, 2009. **69**(18): p. 7385-92.
31. Lu, L.C., et al., *Tumor Heterogeneity in Hepatocellular Carcinoma: Facing the Challenges*. Liver Cancer, 2016. **5**(2): p. 128-38.
32. Friemel, J., et al., *Intratumor heterogeneity in hepatocellular carcinoma*. Clin Cancer Res, 2015. **21**(8): p. 1951-61.
33. Singh, A.K., R. Kumar, and A.K. Pandey, *Hepatocellular Carcinoma: Causes, Mechanism of Progression and Biomarkers*. Current chemical genomics and translational medicine, 2018. **12**: p. 9-26.
34. Tsai, W.L. and R.T. Chung, *Viral hepatocarcinogenesis*. Oncogene, 2010. **29**(16): p. 2309-2324.
35. Petruzzello, A., *Epidemiology of Hepatitis B Virus (HBV) and Hepatitis C Virus (HCV) Related Hepatocellular Carcinoma*. The open virology journal, 2018. **12**: p. 26-32.

36. Benkheil, M., et al., *HCV-induced EGFR-ERK signaling promotes a pro-inflammatory and pro-angiogenic signature contributing to liver cancer pathogenesis*. *Biochemical Pharmacology*, 2018. **155**: p. 305-315.
37. Feld, J., *Update on the Risk of Primary and Recurrent HCC With the Use of DAA Therapy for HCV Infection*. *Gastroenterology & hepatology*, 2019. **15**(6): p. 303-306.
38. Zoller, H. and H. Tilg, *Nonalcoholic fatty liver disease and hepatocellular carcinoma*. *Metabolism*, 2016. **65**(8): p. 1151-1160.
39. DeWeerd, S., *Disease progression : Divergent paths*. *Nature*, 2017.
40. Singal, A.G., A. Pillai, and J. Tiro, *Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis*. *PLoS Med*, 2014. **11**(4): p. e1001624.
41. van Meer, S., et al., *Surveillance for hepatocellular carcinoma is associated with increased survival: Results from a large cohort in the Netherlands*. *J Hepatol*, 2015. **63**(5): p. 1156-63.
42. Singal, A.G., et al., *Detection of hepatocellular carcinoma at advanced stages among patients in the HALT-C trial: where did surveillance fail?* *Am J Gastroenterol*, 2013. **108**(3): p. 425-32.
43. Atiq, O., et al., *An assessment of benefits and harms of hepatocellular carcinoma surveillance in patients with cirrhosis*. *Hepatology*, 2017. **65**(4): p. 1196-1205.
44. Joshi, K., et al., *Hepatocellular carcinoma surveillance: a national survey of current practices in the USA*. *Dig Dis Sci*, 2014. **59**(12): p. 3073-7.
45. Tanaka, H., *Current role of ultrasound in the diagnosis of hepatocellular carcinoma*. *Journal of Medical Ultrasonics*, 2020.
46. Tomasi, T.B., *Structure and Function of Alpha-Fetoprotein*. *Annual Review of Medicine*, 1977. **28**(1): p. 453-465.
47. Johnson, P.J., et al., *The detection of hepatocellular carcinoma using a prospectively developed and validated model based on serological biomarkers*. *Cancer Epidemiol Biomarkers Prev*, 2014. **23**(1): p. 144-53.
48. Aoyagi, Y., *Molecular discrimination between alpha-fetoprotein from patients with hepatocellular-carcinoma and nonneoplastic liver-diseases by their carbohydrate structures (review)*. *International journal of oncology*, 1994. **4**(2): p. 369-83.
49. Aoyagi, Y., *Carbohydrate-based measurements on alpha-fetoprotein in the early diagnosis of hepatocellular carcinoma*. *Glycoconj J*, 1995. **12**(3): p. 194-9.
50. Aoyagi, Y., et al., *Change in fucosylation of alpha-fetoprotein on malignant transformation of liver cells*. *Lancet*, 1986. **1**(8474): p. 210.
51. Block, T., A.S. Mehta, and W.T. London, *Hepatocellular carcinoma of the liver*. *Cancer Biomark*, 2010. **9**(1-6): p. 375-83.
52. Mehta, A. and T.M. Block, *Fucosylated glycoproteins as markers of liver disease*. *Dis Markers*, 2008. **25**(4-5): p. 259-65.
53. Comunale, M.A., et al., *Identification and development of fucosylated glycoproteins as biomarkers of primary hepatocellular carcinoma*. *J Proteome Res*, 2009. **8**(2): p. 595-602.

54. Steel, L.F., et al., *A proteomic approach for the discovery of early detection markers of hepatocellular carcinoma*. *Dis Markers*, 2001. **17**(3): p. 179-89.
55. Singal, A.G., et al., *Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis*. *Cancer Epidemiol Biomarkers Prev*, 2012. **21**(5): p. 793-9.
56. Wang, M., et al., *The Doylestown Algorithm: A Test to Improve the Performance of AFP in the Detection of Hepatocellular Carcinoma*. *Cancer Prev Res (Phila)*, 2016. **9**(2): p. 172-9.
57. Mehta, A.S., et al., *Application of the Doylestown algorithm for the early detection of hepatocellular carcinoma*. *PLoS One*, 2018. **13**(8): p. e0203149.
58. Marrero, J.A. and S. Pelletier, *Hepatocellular carcinoma*. *Clinics in Liver Disease*, 2006. **10**(2): p. 339-51.
59. Lok, A. and B. McMahon, *Chronic hepatitis B*. *Hepatology*, 2001. **34**(6): p. 1225-41.
60. Llovet, J.M., A. Burroughs, and J. Bruix, *Hepatocellular carcinoma*. *Lancet*, 2003. **362**(9399): p. 1907-17.
61. Rudd, P.M., et al., *Glycosylation and the immune system*. *Science*, 2001. **291**(5512): p. 2370-6.
62. Van den Steen, P., et al., *Concepts and principles of O-linked glycosylation*. *Crit Rev Biochem Mol Biol*, 1998. **33**(3): p. 151-208.
63. Opdenakker, G., et al., *Concepts and principles of glycobiology*. *FASEB J*, 1993. **7**(14): p. 1330-7.
64. Reily, C., et al., *Glycosylation in health and disease*. 2019.
65. Christiansen, M.N., et al., *Cell surface protein glycosylation in cancer*. *Proteomics*, 2014. **14**(4-5): p. 525-46.
66. Dennis, J.W., M. Granovsky, and C.E. Warren, *Glycoprotein glycosylation and cancer progression*. *Biochim Biophys Acta*, 1999. **1473**(1): p. 21-34.
67. Silsirivanit, A., *Chapter Five - Glycosylation markers in cancer*, in *Advances in Clinical Chemistry*, G.S. Makowski, Editor. 2019, Elsevier. p. 189-213.
68. Pinho, S.S. and C.A. Reis, *Glycosylation in cancer: Mechanisms and clinical implications*. 2015.
69. Julien, S., et al., *How Do Gangliosides Regulate RTKs Signaling?* *Cells*, 2013. **2**(4): p. 751-767.
70. Akella, N.M., L. Ciraku, and M.J. Reginato, *Fueling the fire: emerging role of the hexosamine biosynthetic pathway in cancer*. *BMC biology*, 2019. **17**(1): p. 52-52.
71. Chiaradonna, F., F. Ricciardiello, and R. Palorini, *The Nutrient-Sensing Hexosamine Biosynthetic Pathway as the Hub of Cancer Metabolic Rewiring*. *Cells*, 2018. **7**(6): p. 53.
72. Moremen, K.W., M. Tiemeyer, and A.V. Nairn, *Vertebrate protein glycosylation: diversity, synthesis and function*. *Nature reviews. Molecular cell biology*, 2012. **13**(7): p. 448-462.



73. Roger, K.B., *Structure, Expression, and Regulation of UDP-GlcNAc: Dolichol Phosphate GlcNAc-1-Phosphate Transferase (DPAGT1)*. *Current Drug Targets*, 2009. **10**(6): p. 477-482.
74. Bieberich, E., *Synthesis, Processing, and Function of N-glycans in N-glycoproteins*. *Advances in neurobiology*, 2014. **9**: p. 47-70.
75. Kornfeld, R. and S. Kornfeld, *ASSEMBLY OF ASPARAGINE-LINKED OLIGOSACCHARIDES*. *Annual Review of Biochemistry*, 1985. **54**(1): p. 631-664.
76. Drake, R.R., et al., *MALDI Mass Spectrometry Imaging of N-Linked Glycans in Tissues BT - Glycobiophysics*, Y. Yamaguchi and K. Kato, Editors. 2018, Springer Singapore: Singapore. p. 59-76.
77. Bhide, G.P. and K.J. Colley, *Sialylation of N-glycans: mechanism, cellular compartmentalization and function*. *Histochemistry and Cell Biology*, 2017. **147**(2): p. 149-174.
78. Chang, I.J., M. He, and C.T. Lam, *Congenital disorders of glycosylation*. *Annals of translational medicine*, 2018. **6**(24): p. 477-477.
79. Christiansen, M.N., et al., *Cell surface protein glycosylation in cancer*. *PROTEOMICS*, 2014. **14**(4-5): p. 525-546.
80. Varki, A., *Biological roles of glycans*. *Glycobiology*, 2017. **27**(1).
81. Cherepanova, N., S. Shrimal, and R. Gilmore, *N-linked glycosylation and homeostasis of the endoplasmic reticulum*. *Current opinion in cell biology*, 2016. **41**: p. 57-65.
82. Helenius, A., Aebi, and Markus, *Intracellular Functions of N-Linked Glycans*. *Science*, 2001. **291**(5512): p. 2364.
83. Gilbert, H.J. and N. Chandra, *Editorial overview: Carbohydrate–protein interactions and glycosylation: integrating structural biology, informatics and systems modelling to understand glycan structure and glycan-protein interactions*. *Current Opinion in Structural Biology*, 2016. **40**: p. v-viii.
84. Lauc, G., et al., *Mechanisms of disease: The human N-glycome*. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 2016. **1860**(8): p. 1574-1582.
85. Ohtsubo, K. and J.D. Marth, *Glycosylation in Cellular Mechanisms of Health and Disease*. *Cell*, 2006. **126**(5): p. 855-867.
86. Stowell, S.R., T. Ju, and R.D. Cummings, *Protein glycosylation in cancer*. *Annual review of pathology*, 2015. **10**: p. 473-510.
87. Hakomori, S.I. and W.T. Murakami, *Glycolipids of hamster fibroblasts and derived malignant-transformed cell lines*. *Proceedings of the National Academy of Sciences of the United States of America*, 1968. **59**(1): p. 254-261.
88. Munkley, J. and D.J. Elliott, *Hallmarks of glycosylation in cancer*. *Oncotarget*, 2016. **7**(23).
89. Lau, K.S., et al., *Complex N-Glycan Number and Degree of Branching Cooperate to Regulate Cell Proliferation and Differentiation*. *Cell*, 2007. **129**(1): p. 123-134.
90. English, N.M., J.F. Lesley, and R. Hyman, *Site-specific De-N-glycosylation of CD44 Can Activate Hyaluronan Binding, and CD44 Activation States*

- Show Distinct Threshold Densities for Hyaluronan Binding.* Cancer Research, 1998. **58**(16): p. 3736.
91. Paszek, M.J., et al., *The cancer glycocalyx mechanically primes integrin-mediated growth and survival.* Nature, 2014. **511**(7509): p. 319-325.
  92. Taniguchi, N. and Y. Kizuka, *Glycans and cancer: Role of N-Glycans in cancer biomarker, progression and metastasis, and therapeutics.* Advances in Cancer Research, 2015. **126**.
  93. Cheung, P. and J.W. Dennis, *Mgat5 and Pten interact to regulate cell growth and polarity.* Glycobiology, 2007. **17**(7): p. 767-773.
  94. Lagana, A., et al., *Galectin Binding to Mgat5-Modified N-Glycans Regulates Fibronectin Matrix Remodeling in Tumor Cells Galectin Binding to Mgat5-Modified N-Glycans Regulates Fibronectin Matrix Remodeling in Tumor Cells.* 2006. **26**(8): p. 3181-3193.
  95. Oliveira-Ferrer, L., K. Legler, and K. Milde-Langosch, *Role of protein glycosylation in cancer metastasis.* Seminars in Cancer Biology, 2017. **44**: p. 141-152.
  96. Vajaria, B.N. and P.S. Patel, *Glycosylation: a hallmark of cancer?* Glycoconjugate Journal, 2017. **34**(2): p. 147-156.
  97. Rodrigues, J.G., et al., *Glycosylation in cancer: Selected roles in tumour progression, immune modulation and metastasis.* Cellular Immunology, 2018. **333**: p. 46-57.
  98. Dube, D.H. and C.R. Bertozzi, *Glycans in cancer and inflammation — potential for therapeutics and diagnostics.* Nature Reviews Drug Discovery, 2005. **4**(6): p. 477-488.
  99. Agrawal, P., et al., *A Systems Biology Approach Identifies FUT8 as a Driver of Melanoma Metastasis.* Cancer cell, 2017. **31**(6): p. 804-819.e7.
  100. Chen, C.-Y., et al., *Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(2): p. 630-635.
  101. Wang, Y., et al., *Loss of  $\alpha$ 1,6-fucosyltransferase inhibits chemical-induced hepatocellular carcinoma and tumorigenesis by down-regulating several cell signaling pathways.* The FASEB Journal, 2015. **29**(8): p. 3217-3227.
  102. Scott, E. and J. Munkley, *Glycans as Biomarkers in Prostate Cancer.* International journal of molecular sciences, 2019. **20**(6): p. 1389.
  103. Norton, P.A. and A.S. Mehta, *Expression of genes that control core fucosylation in hepatocellular carcinoma: Systematic review.* World journal of gastroenterology, 2019. **25**(23): p. 2947-2960.
  104. Miyoshi, E., et al., *The  $\alpha$ 1-6-fucosyltransferase gene and its biological significance.* Biochimica et Biophysica Acta (BBA) - General Subjects, 1999. **1473**(1): p. 9-20.
  105. Treanor, B., *B-cell receptor: from resting state to activate.* Immunology, 2012. **136**(1): p. 21-27.
  106. Li, W., et al., *Core fucosylation of  $\mu$  heavy chains regulates assembly and intracellular signaling of precursor B cell receptors.* The Journal of biological chemistry, 2012. **287**(4): p. 2500-2508.

107. Wang, X., et al., *Phenotype Changes of Fut8 Knockout Mouse: Core Fucosylation Is Crucial for the Function of Growth Factor Receptor(s)*, in *Methods in Enzymology*. 2006, Academic Press. p. 11-22.
108. Okeley, N.M., et al., *Development of orally active inhibitors of protein and cellular fucosylation*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(14): p. 5404-5409.
109. Zhou, Y., et al., *Inhibition of fucosylation by 2-fluorofucose suppresses human liver cancer HepG2 cell proliferation and migration as well as tumor formation*. Scientific reports, 2017. **7**(1): p. 11563-11563.
110. Lübbers, J., E. Rodríguez, and Y. van Kooyk, *Modulation of Immune Tolerance via Siglec-Sialic Acid Interactions*. Frontiers in immunology, 2018. **9**: p. 2807-2807.
111. Stencel-Baerenwald, J.E., et al., *The sweet spot: defining virus-sialic acid interactions*. Nature reviews. Microbiology, 2014. **12**(11): p. 739-749.
112. Tortorici, M.A., et al., *Structural basis for human coronavirus attachment to sialic acid receptors*. Nature structural & molecular biology, 2019. **26**(6): p. 481-489.
113. Matrosovich, M., G. Herrler, and H.D. Klenk, *Sialic Acid Receptors of Viruses*, in *SialoGlyco Chemistry and Biology II: Tools and Techniques to Identify and Capture Sialoglycans*, R. Gerardy-Schahn, P. Delannoy, and M. von Itzstein, Editors. 2015, Springer International Publishing: Cham. p. 1-28.
114. Läubli, H. and L. Borsig, *Selectins promote tumor metastasis*. Seminars in Cancer Biology, 2010. **20**(3): p. 169-177.
115. Seales, E.C., et al., *Ras oncogene directs expression of a differentially sialylated, functionally altered  $\beta$ 1 integrin*. Oncogene, 2003. **22**(46): p. 7137-7145.
116. Elola, M.T., et al., *Assembly, organization and regulation of cell-surface receptors by lectin-glycan complexes*. Biochem J, 2015. **469**(1): p. 1-16.
117. Cheung, P. and J.W. Dennis, *Mgat5 and Pten interact to regulate cell growth and polarity*. Glycobiology, 2007. **17**(7): p. 767-73.
118. Takahashi, M., et al., *Disease-associated glycans on cell surface proteins*. Mol Aspects Med, 2016. **51**: p. 56-70.
119. Nishimori, I., et al., *N-acetylgalactosamine glycosylation of MUC1 tandem repeat peptides by pancreatic tumor cell extracts*. Cancer Res, 1994. **54**(14): p. 3738-3744.
120. Hollingsworth, M.A., et al., *Expression of MUC1, MUC2, MUC3 and MUC4 mucin mRNAs in human pancreatic and intestinal tumor cell lines*. Int J Cancer, 1994. **57**(2): p. 198-203.
121. Chambers, J.A., et al., *Developmental expression of mucin genes MUC1 and MUC2*. J Cell Sci, 1994. **107** ( Pt 2): p. 413-424.
122. Liu, Y., et al., *Identification and confirmation of biomarkers using an integrated platform for quantitative analysis of glycoproteins and their glycosylations*. J Proteome Res, 2010. **9**(2): p. 798-805.
123. Hollingsworth, M.A. and B.J. Swanson, *Mucins in cancer: protection and control of the cell surface*. Nat Rev Cancer, 2004. **4**(1): p. 45-60.

124. Batra, S.K., et al., *Transfection of the human Muc 1 mucin gene into a poorly differentiated human pancreatic tumor cell line, Panc1: integration, expression and ultrastructural changes*. J Cell Sci, 1991. **100** ( Pt 4): p. 841-849.
125. Andrianifahanana, M., et al., *Mucin (MUC) gene expression in human pancreatic adenocarcinoma and chronic pancreatitis: a potential role of MUC4 as a tumor marker of diagnostic significance*. Clin Cancer Res, 2001. **7**(12): p. 4033-4040.
126. Shibahara, H., et al., *MUC4 is a novel prognostic factor of intrahepatic cholangiocarcinoma-mass forming type*. Hepatology, 2004. **39**(1): p. 220-229.
127. Andrianifahanana, M., et al., *Synergistic induction of the MUC4 mucin gene by interferon-gamma and retinoic acid in human pancreatic tumour cells involves a reprogramming of signalling pathways*. Oncogene, 2005. **24**(40): p. 6143-6154.
128. Singh, A.P., et al., *Aberrant expression of transmembrane mucins, MUC1 and MUC4, in human prostate carcinomas*. Prostate, 2006. **66**(4): p. 421-429.
129. Ludwig, J.A. and J.N. Weinstein, *Biomarkers in cancer staging, prognosis and treatment selection*. Nat Rev Cancer, 2005. **5**(11): p. 845-856.
130. Springer, S.A. and P. Gagneux, *Glycomics: revealing the dynamic ecology and evolution of sugar molecules*. J Proteomics, 2016. **135**: p. 90-100.
131. Springer, S.A. and P. Gagneux, *Glycan evolution in response to collaboration, conflict, and constraint*. J Biol Chem, 2013. **288**(10): p. 6904-11.
132. Dennis, J.W., I.R. Nabi, and M. Demetriou, *Metabolism, cell surface organization, and disease*. Cell, 2009. **139**(7): p. 1229-41.
133. Goode, G., et al., *MUC1 facilitates metabolomic reprogramming in triple-negative breast cancer*. PLoS One, 2017. **12**(5): p. e0176820.
134. Lai, K.K., D. Kolippakkam, and L. Beretta, *Comprehensive and quantitative proteome profiling of the mouse liver and plasma*. Hepatology, 2008. **47**(3): p. 1043-1051.
135. Mato, J.M., F. He, and L. Beretta, *The 2006 Human Liver Proteome Project (HLPP) Workshops*. Proteomics Clin Appl, 2007. **1**(5): p. 442-445.
136. Beretta, L., *Liver proteomics applied to translational research in liver disease and cancer*. Proteomics Clin Appl, 2010. **4**(4): p. 359-361.
137. Yi, X., et al., *Association of mortalin (HSPA9) with liver cancer metastasis and prediction for early tumor recurrence*. Mol Cell Proteomics, 2008. **7**(2): p. 315-325.
138. Goldman, R., et al., *Detection of hepatocellular carcinoma using glycomic analysis*. Clin Cancer Res, 2009. **15**(5): p. 1808-1813.
139. Lattova, E., et al., *Mass spectrometric study of N-glycans from serum of woodchucks with liver cancer*. Rapid Commun Mass Spectrom, 2009. **23**(18): p. 2983-2995.

140. Isailovic, D., et al., *Profiling of human serum glycans associated with liver cancer and cirrhosis by IMS-MS*. J Proteome Res, 2008. **7**(3): p. 1109-1117.
141. Amano, M. and S. Nishimura, *Large-scale glycomics for discovering cancer-associated N-glycans by integrating glycoblotting and mass spectrometry*. Methods Enzymol, 2010. **478**: p. 109-125.
142. Fang, M., et al., *Serum N-glycome biomarker for monitoring development of DENA-induced hepatocellular carcinoma in rat*. Mol Cancer, 2010. **9**: p. 215-215.
143. Nakagawa, T., et al., *Glycomic analysis of alpha-fetoprotein L3 in hepatoma cell lines and hepatocellular carcinoma patients*. J Proteome Res, 2008. **7**(6): p. 2222-2233.
144. Nakagawa, T., et al., *Glycomic analyses of glycoproteins in bile and serum during rat hepatocarcinogenesis*. J Proteome Res, 2010. **9**(10): p. 4888-4896.
145. An, H.J. and C.B. Lebrilla, *A glycomics approach to the discovery of potential cancer biomarkers*. Methods Mol Biol, 2010. **600**: p. 199-213.
146. An, H.J., et al., *Glycomics and disease markers*. Curr Opin Chem Biol, 2009. **13**(5-6): p. 601-607.
147. Packer, N.H., et al., *Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11-13, 2006)*. Proteomics, 2008. **8**(1): p. 8-20.
148. Kirmiz, C., et al., *A serum glycomics approach to breast cancer biomarkers*. Mol Cell Proteomics, 2007. **6**(1): p. 43-55.
149. Zhu, J., et al., *Analysis of serum haptoglobin fucosylation in hepatocellular carcinoma and liver cirrhosis of different etiologies*. Journal of Proteome Research, 2014. **13**(6): p. 2986-2997.
150. Bernardi, C., et al., *Effects of MicroRNAs on Fucosyltransferase 8 (FUT8) Expression in Hepatocarcinoma Cells*. PLoS ONE, 2013.
151. Block, T.M., et al., *Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(3): p. 779-84.
152. Comunale, M.A., et al., *Proteomic Analysis of Serum Associated Fucosylated Glycoproteins in the Development of Primary Hepatocellular Carcinoma*.
153. Comunale, M.A., et al., *Linkage specific fucosylation of alpha-1-antitrypsin in liver cirrhosis and cancer patients: implications for a biomarker of hepatocellular carcinoma*. PloS one, 2010. **5**(8): p. e12419-e12419.
154. Ma, J., et al., *Quantitative analysis of core fucosylation of serum proteins in liver diseases by LC-MS-MRM*. Journal of proteomics, 2018. **189**: p. 67-74.
155. Wang, M., et al., *Changes in the Glycosylation of Kininogen and the Development of a Kininogen-Based Algorithm for the Early Detection of HCC*. Cancer Epidemiol Biomarkers Prev, 2017. **26**(5): p. 795-803.

156. Wang, Y., et al., *Loss of alpha1,6-fucosyltransferase suppressed liver regeneration: implication of core fucose in the regulation of growth factor receptor-mediated cellular signaling*. Sci Rep, 2015. **5**: p. 8264.
157. Mehta, A., H. Herrera, and T. Block, *Glycosylation and liver cancer*. Advances in cancer research, 2015. **126**: p. 257-279.
158. Kamada, Y., et al., *Serum Fucosylated Haptoglobin as a Novel Diagnostic Biomarker for Predicting Hepatocyte Ballooning and Nonalcoholic Steatohepatitis*. PloS one, 2013. **8**(6): p. e66328-e66328.
159. Kizuka, Y. and N. Taniguchi, *Enzymes for N-Glycan branching and their genetic and nongenetic regulation in cancer*. 2016.
160. Mehta, A., et al., *Increased levels of tetra-antennary N-linked glycan but not core fucosylation are associated with hepatocellular carcinoma tissue*. Cancer Epidemiology Biomarkers and Prevention, 2012. **21**(6): p. 925-933.
161. West, C.A., et al., *N-Linked Glycan Branching and Fucosylation Are Increased Directly in Hcc Tissue As Determined through in Situ Glycan Imaging*. Journal of Proteome Research, 2018. **17**(10): p. 3454-3462.
162. Cui, J., et al., *N-glycosylation by N-acetylglucosaminyltransferase V enhances the interaction of CD147/basigin with integrin  $\beta$ 1 and promotes HCC metastasis*. The Journal of pathology, 2018. **245**(1): p. 41-52.
163. Cao, Y., et al., *Expression of MUC1, Thomsen-Friedenreich antigen, Tn, sialosyl-Tn, and  $\alpha$ 2,6-linked sialic acid in hepatocellular carcinomas and preneoplastic hepatocellular lesions*. Virchows Archiv, 1999. **434**(6): p. 503-509.
164. Chen, X., et al., *ST6Gal-I modulates docetaxel sensitivity in human hepatocarcinoma cells via the p38 MAPK/caspase pathway*. Oncotarget, 2016. **7**(32): p. 51955-51964.
165. Gruszewska, E., et al., *Total and free serum sialic acid concentration in liver diseases*. BioMed research international, 2014. **2014**: p. 876096-876096.
166. Kongtawelert, P., et al., *Role of serum total sialic acid in differentiating cholangiocarcinoma from hepatocellular carcinoma*. World journal of gastroenterology, 2003. **9**(10): p. 2178-2181.
167. Mondal, G., et al., *Alterations of glycan branching and differential expression of sialic acid on alpha fetoprotein among hepatitis patients*. Glycoconjugate Journal, 2011. **28**(1): p. 1-9.
168. Powers, T.W., et al., *Two-Dimensional N-Glycan Distribution Mapping of Hepatocellular Carcinoma Tissues by MALDI-Imaging Mass Spectrometry*. Biomolecules, 2015. **5**: p. 2554-2572.
169. Kwan-Shuen Chan, K., et al., *Secretory Stanniocalcin 1 promotes metastasis of hepatocellular carcinoma through activation of JNK signaling pathway*. Cancer Lett, 2017.
170. Liu, C., et al., *c-Jun-dependent  $\beta$ 3GnT8 promotes tumorigenesis and metastasis of hepatocellular carcinoma by inducing CD147 glycosylation and altering N-glycan patterns*. Oncotarget, 2018. **9**(26): p. 18327-18340.

171. Chen, R., et al., *Development of a combined chemical and enzymatic approach for the mass spectrometric identification and quantification of aberrant N-glycosylation*. Journal of proteomics, 2012. **75**(5): p. 1666-74.
172. Zhao, Y., et al., *Fragmentation and site-specific quantification of core fucosylated glycoprotein by multiple reaction monitoring-mass spectrometry*. Analytical Chemistry, 2011. **83**(22): p. 8802-9.
173. Ahn, Y.H., et al., *A lectin-coupled, multiple reaction monitoring based quantitative analysis of human plasma glycoproteins by mass spectrometry*. Analytical and bioanalytical chemistry, 2012. **402**(6): p. 2101-12.
174. Ahn, Y.H., et al., *A lectin-coupled, targeted proteomic mass spectrometry (MRM MS) platform for identification of multiple liver cancer biomarkers in human plasma*. Journal of proteomics, 2012.
175. Kamiyama, T., et al., *Identification of novel serum biomarkers of hepatocellular carcinoma using glycomic analysis*. Hepatology, 2013. **57**(6): p. 2314-25.
176. Liu, X.E., et al., *N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus*. Hepatology, 2007. **46**(5): p. 1426-35.
177. Ajdukiewicz, A.B., et al., *Alpha-fetoprotein glycosylation is abnormal in some hepatocellular carcinoma, including white patients with a normal alpha-fetoprotein concentration*. Cancer letters, 1993. **74**(1-2): p. 43-50.
178. Amano, M. and S. Nishimura, *Large-scale glycomics for discovering cancer-associated N-glycans by integrating glycoblotting and mass spectrometry*. Methods in enzymology, 2010. **478**: p. 109-25.
179. Aoyagi, Y., et al., *Highly enhanced fucosylation of alpha-fetoprotein in patients with germ cell tumor*. Cancer, 1993. **72**(2): p. 615-8.
180. Comunale, M.A., et al., *Proteomic analysis of serum associated fucosylated glycoproteins in the development of primary hepatocellular carcinoma*. Journal of Proteome Research, 2006. **5**(2): p. 308-15.
181. Comunale, M.A., et al., *Linkage specific fucosylation of alpha-1-antitrypsin in liver cirrhosis and cancer patients: implications for a biomarker of hepatocellular carcinoma*. PLoS One, 2010. **5**(8): p. e12419.
182. Dai, Z., et al., *Identification and analysis of altered alpha1,6-fucosylated glycoproteins associated with hepatocellular carcinoma metastasis*. Proteomics, 2006. **6**(21): p. 5857-67.
183. Dai, Z., et al., *Lectin-based glycoproteomics to explore and analyze hepatocellular carcinoma-related glycoprotein markers*. Electrophoresis, 2009. **30**(17): p. 2957-66.
184. Debruyne, E.N., et al., *Diagnostic value of the hemopexin N-glycan profile in hepatocellular carcinoma patients*. Clin Chem, 2010. **56**(5): p. 823-31.
185. Kelleher, P.C., et al., *Altered glycosylation of alpha-fetoprotein in hepadnavirus-induced hepatocellular carcinoma of the woodchuck*. Cancer Lett, 1992. **63**(2): p. 93-9.

186. Wang, Y., et al., *Loss of alpha1,6-fucosyltransferase inhibits chemical-induced hepatocellular carcinoma and tumorigenesis by down-regulating several cell signaling pathways*. FASEB J, 2015. **29**(8): p. 3217-27.
187. Yin, H., et al., *Mass-selected site-specific core-fucosylation of ceruloplasmin in alcohol-related hepatocellular carcinoma*. J Proteome Res, 2014. **13**(6): p. 2887-96.
188. Zhang, Y., et al., *ESI-LC-MS Method for Haptoglobin Fucosylation Analysis in Hepatocellular Carcinoma and Liver Cirrhosis*. J Proteome Res, 2015. **14**(12): p. 5388-95.
189. Zhu, J., et al., *Aberrant fucosylation of glycosphingolipids in human hepatocellular carcinoma tissues*. Liver Int, 2014. **34**(1): p. 147-60.
190. Zhu, J., et al., *Mass Spectrometric N-Glycan Analysis of Haptoglobin from Patient Serum Samples Using a 96-Well Plate Format*. J Proteome Res, 2015. **14**(11): p. 4932-9.
191. Comunale, M.A., et al., *Total serum glycan analysis is superior to lectin-FLISA for the early detection of hepatocellular carcinoma*. Proteomics. Clinical applications, 2013. **7**: p. 690-700.
192. Chen, C.Y., et al., *Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer*. Proc Natl Acad Sci U S A, 2013. **110**(2): p. 630-5.
193. Romano, P.R., et al., *Development of recombinant Aleuria aurantia lectins with altered binding specificities to fucosylated glycans*. Biochem Biophys Res Commun, 2011. **414**(1): p. 84-9.
194. Liu, Y., et al., *beta1,6-N-acetylglucosaminyltransferase V predicts recurrence and survival of patients with clear-cell renal cell carcinoma after surgical resection*. World J Urol, 2015. **33**(11): p. 1791-9.
195. Powers, T.W., et al., *Two-Dimensional N-Glycan Distribution Mapping of Hepatocellular Carcinoma Tissues by MALDI-Imaging Mass Spectrometry*. Biomolecules, 2015. **5**(4): p. 2554-2572.
196. West, C.A., et al., *N-Linked Glycan Branching and Fucosylation Are Increased Directly in Hcc Tissue As Determined through in Situ Glycan Imaging*. J Proteome Res, 2018. **17**(10): p. 3454-3462.
197. Mehta, A., et al., *Increased levels of tetra-antennary N-linked glycan but not core fucosylation are associated with hepatocellular carcinoma tissue*. Cancer Epidemiol Biomarkers Prev, 2012. **21**(6): p. 925-33.
198. Moriwaki, K., et al., *A high expression of GDP-fucose transporter in hepatocellular carcinoma is a key factor for increases in fucosylation*. Glycobiology, 2007. **17**(12): p. 1311-20.
199. Mehta, A., et al., *Intrinsic hepatocyte dedifferentiation is accompanied by upregulation of mesenchymal markers, protein sialylation and core alpha 1,6 linked fucosylation*. Sci Rep, 2016. **6**: p. 27965.
200. Chen, C.Y., et al., *Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(2): p. 630-5.
201. Mehta, A., et al., *Increased Levels of Tetra-antennary N-Linked Glycan but Not Core Fucosylation Are Associated with Hepatocellular Carcinoma*



- Tissue. Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2012. **21**(6): p. 925-933.
202. Chaurand, P., M. Stoeckli, and R.M. Caprioli, *Direct profiling of proteins in biological tissue sections by MALDI mass spectrometry*. Analytical Chemistry, 1999. **71**(23): p. 5263-70.
  203. Caldwell, R.L. and R.M. Caprioli, *Tissue profiling by mass spectrometry: a review of methodology and applications*. Molecular & cellular proteomics : MCP, 2005. **4**(4): p. 394-401.
  204. Schwamborn, K. and R.M. Caprioli, *MALDI imaging mass spectrometry--painting molecular pictures*. Molecular oncology, 2010. **4**(6): p. 529-38.
  205. Schwamborn, K. and R.M. Caprioli, *Molecular imaging by mass spectrometry--looking beyond classical histology*. Nature reviews. Cancer, 2010. **10**(9): p. 639-46.
  206. Dai, C., et al., *Using boronolactin in MALDI-MS imaging for the histological analysis of cancer tissue expressing the sialyl Lewis X antigen*. Chemical communications, 2011. **47**(37): p. 10338-40.
  207. Kim, J., et al., *Elevated plasma osteopontin levels in patients with hepatocellular carcinoma*. Am J Gastroenterol, 2006. **101**(9): p. 2051-9.
  208. Deng, B., et al., *Correlation and prognostic value of osteopontin and Bcl-2 in hepatocellular carcinoma patients after curative resection*. Oncol Rep, 2013. **30**(6): p. 2795-803.
  209. Abu El Makarem, M.A., et al., *Diagnostic significance of plasma osteopontin in hepatitis C virus-related hepatocellular carcinoma*. Ann Hepatol, 2011. **10**(3): p. 296-305.
  210. Qiao, B., et al., *Detection and identification of peroxiredoxin 3 as a biomarker in hepatocellular carcinoma by a proteomic approach*. Int J Mol Med, 2012. **29**(5): p. 832-40.
  211. Huang, Y., et al., *LC-MS/MS isomeric profiling of permethylated N-glycans derived from serum haptoglobin of hepatocellular carcinoma (HCC) and cirrhotic patients*. Electrophoresis, 2017. **38**(17): p. 2160-2167.
  212. Zhang, Y., et al., *ESI-LC-MS Method for Haptoglobin Fucosylation Analysis in Hepatocellular Carcinoma and Liver Cirrhosis*. Journal of proteome research, 2015. **14**(12): p. 5388-5395.
  213. Zhu, J., et al., *Differential Quantitative Determination of Site-Specific Intact N-Glycopeptides in Serum Haptoglobin between Hepatocellular Carcinoma and Cirrhosis Using LC-ETHcD-MS/MS*. Journal of proteome research, 2019. **18**(1): p. 359-371.
  214. Darebna, P., et al., *Changes in the expression of N- and O-glycopeptides in patients with colorectal cancer and hepatocellular carcinoma quantified by full-MS scan FT-ICR and multiple reaction monitoring*. Journal of proteomics, 2017. **153**: p. 44-52.
  215. Goldman, R. and M. Sanda, *Targeted methods for quantitative analysis of protein glycosylation*. Proteomics. Clinical applications, 2015. **9**(1-2): p. 17-32.

216. Resson, H.W., et al., *Analysis of MALDI-TOF mass spectrometry data for discovery of peptide and glycan biomarkers of hepatocellular carcinoma*. Journal of proteome research, 2008. **7**(2): p. 603-610.
217. Biosoft, P. *Mass Spectrometry*. [cited 2020 March 23]; Available from: [http://www.premierbiosoft.com/tech\\_notes/mass-spectrometry.html](http://www.premierbiosoft.com/tech_notes/mass-spectrometry.html).
218. Chong, Y.-K., et al., *Clinical Mass Spectrometry in the Bioinformatics Era: A Hitchhiker's Guide*. Computational and Structural Biotechnology Journal, 2018. **16**: p. 316-334.
219. Savaryn, J.P., T.K. Toby, and N.L. Kelleher, *A researcher's guide to mass spectrometry-based proteomics*. Proteomics, 2016. **16**(18): p. 2435-2443.
220. Haag, A.M., *Mass Analyzers and Mass Spectrometers*, in *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*, H. Mirzaei and M. Carrasco, Editors. 2016, Springer International Publishing: Cham. p. 157-169.
221. Caprioli, R.M., *Imaging Mass Spectrometry: A Perspective*. Journal of biomolecular techniques : JBT, 2019. **30**(1): p. 7-11.
222. Dekker, J.P. and J.A. Branda, *MALDI-TOF Mass Spectrometry in the Clinical Microbiology Laboratory*. Clinical Microbiology Newsletter, 2011. **33**(12): p. 87-93.
223. Murayama, C., Y. Kimura, and M. Setou, *Imaging mass spectrometry: principle and application*. Biophysical reviews, 2009. **1**(3): p. 131-131.
224. Ly, A., et al., *Site-to-Site Reproducibility and Spatial Resolution in MALDI-MSI of Peptides from Formalin-Fixed Paraffin-Embedded Samples*. Proteomics. Clinical applications, 2019. **13**(1): p. e1800029-e1800029.
225. Pitt, J.J., *Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry*. The Clinical biochemist. Reviews, 2009. **30**(1): p. 19-34.
226. Li, S., et al., *Selective fragmentation of the N-glycan moiety and protein backbone of ribonuclease B on an Orbitrap Fusion Lumos Tribrid mass spectrometer*. Rapid Communications in Mass Spectrometry, 2018. **32**(23): p. 2031-2039.
227. Saba, J., et al., *Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation*. International journal of proteomics, 2012. **2012**: p. 560391-560391.
228. Singh, C., et al., *Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins*. Journal of Proteome Research, 2012. **11**(9): p. 4517-4525.
229. Blackburn, J.W.T., et al., *Laser Desorption/Ionization Coupled to FTICR Mass Spectrometry for Studies of Natural Organic Matter*. Anal Chem, 2017. **89**(8): p. 4382-4386.
230. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: a primer*. Mass Spectrom Rev, 1998. **17**(1): p. 1-35.
231. *Bruker Solarix FT-ICR Manual*, B.D. GmbH, Editor.

232. Jurinke, C., P. Oeth, and D. van den Boom, *MALDI-TOF mass spectrometry*. Molecular Biotechnology, 2004. **26**(2): p. 147-163.
233. Miki, A., et al., *MALDI-TOF and MALDI-FTICR imaging mass spectrometry of methamphetamine incorporated into hair*. Journal of Mass Spectrometry, 2011. **46**(4): p. 411-416.
234. *Bruker RapifleX MALDI-TOF/TOF Brochure*, B.D. GmbH, Editor.
235. *Thermo Scientific Orbitrap Fusion Lumos Tribrid Mass Spectrometer Brochure*, T. Scientific, Editor.
236. Hu, Q., et al., *The Orbitrap: a new mass spectrometer*. J Mass Spectrom, 2005. **40**(4): p. 430-43.
237. Proteomics, C. *Orbitrap Fusion Lumos Tribrid Mass Spectrometer*. [cited 2020 March 23].
238. Kuzmanov, U., H. Kosanam, and E.P. Diamandis, *The sweet and sour of serological glycoprotein tumor biomarker quantification*. BMC medicine, 2013. **11**: p. 31-31.
239. Villanueva, A., et al., *Genomics and Signaling Pathways in Hepatocellular Carcinoma*.
240. Powers, T.W., et al., *A MALDI Imaging Mass Spectrometry Workflow for Spatial Profiling Analysis of N-linked Glycan Expression in Tissues*. Analytical Chemistry, 2013. **85**(20): p. 9799-9806.
241. Reiding, K.R., et al., *High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification*. Anal Chem, 2014. **86**(12): p. 5784-93.
242. Miyoshi, E., et al., *The alpha1-6-fucosyltransferase gene and its biological significance*. Biochim Biophys Acta, 1999. **1473**(1): p. 9-20.
243. Korczak, B., et al., *Branching N-linked oligosaccharides in breast cancer*. Adv Exp Med Biol, 1994. **353**: p. 95-104.
244. Dennis, J.W., *N-linked oligosaccharide processing and tumor cell biology*. Semin Cancer Biol, 1991. **2**(6): p. 411-20.
245. Bruyneel, E.A., et al., *Altered glycosylation in Madin-Darby canine kidney (MDCK) cells after transformation by murine sarcoma virus*. Clin Exp Metastasis, 1990. **8**(3): p. 241-53.
246. Hiraizumi, S., et al., *Comparative study of the N-linked oligosaccharides released from normal human esophageal epithelium and esophageal squamous carcinoma*. Jpn J Cancer Res, 1990. **81**(4): p. 363-71.
247. Hass, R., et al., *Alterations in glycosylation and lectin pattern during phorbol ester-induced differentiation of U937 cells*. Cancer Res, 1990. **50**(2): p. 323-7.
248. Capony, F., et al., *Increased secretion, altered processing, and glycosylation of pro-cathepsin D in human mammary cancer cells*. Cancer Res, 1989. **49**(14): p. 3904-9.
249. Hubbard, S.C., *Regulation of glycosylation. The influence of protein structure on N-linked oligosaccharide processing*. J Biol Chem, 1988. **263**(36): p. 19303-17.

250. Bolscher, J.G., et al., *Ras (proto)oncogene induces N-linked carbohydrate modification: temporal relationship with induction of invasive potential*. EMBO J, 1988. **7**(11): p. 3361-8.
251. Hubbard, S.C., *Differential effects of oncogenic transformation on N-linked oligosaccharide processing at individual glycosylation sites of viral glycoproteins*. J Biol Chem, 1987. **262**(34): p. 16403-11.
252. Norton, P.A., et al., *N-linked glycosylation of the liver cancer biomarker GP73*. J Cell Biochem, 2008. **104**(1): p. 136-49.
253. Comunale, M.A., et al., *Identification and development of fucosylated glycoproteins as biomarkers of primary hepatocellular carcinoma*. Journal of Proteome Research, 2009. **8**(2): p. 595-602.
254. Wang, M., et al., *Novel fucosylated biomarkers for the early detection of hepatocellular carcinoma*. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2009. **18**(6): p. 1914-21.
255. Hann, H.W., et al., *Analysis of GP73 in patients with HCC as a function of anti-cancer treatment*. Cancer biomarkers : section A of Disease markers, 2010. **7**(6): p. 269-73.
256. Comunale, M.A., et al., *Novel changes in glycosylation of serum Apo-J in patients with hepatocellular carcinoma*. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2011. **20**(6): p. 1222-9.
257. Wang, M., et al., *Changes in the glycosylation of kininogen and the development of a kininogen based algorithm for the early detection of HCC*. Cancer Epidemiol Biomarkers Prev, 2017.
258. Powers, T.W., et al., *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): p. 9799-9806.
259. Powers, T.W., et al., *MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays*. PLoS One, 2014. **9**(9): p. e106255.
260. Drake, R.R., et al., *MALDI Mass Spectrometry Imaging of N-Linked Glycans in Cancer Tissues*. Adv Cancer Res, 2017. **134**: p. 85-116.
261. Heijs, B., et al., *Multimodal Mass Spectrometry Imaging of N-Glycans and Proteins from the Same Tissue Section*. Anal Chem, 2016. **88**(15): p. 7745-53.
262. Holst, S., et al., *Linkage-Specific in Situ Sialic Acid Derivatization for N-Glycan Mass Spectrometry Imaging of Formalin-Fixed Paraffin-Embedded Tissues*. Anal Chem, 2016. **88**(11): p. 5904-13.
263. Everest-Dass, A.V., et al., *N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues*. Mol Cell Proteomics, 2016. **15**(9): p. 3003-16.

264. Briggs, M.T., et al., *MALDI mass spectrometry imaging of N-glycans on tibial cartilage and subchondral bone proteins in knee osteoarthritis*. Proteomics, 2016. **16**(11-12): p. 1736-41.
265. Gustafsson, O.J., et al., *MALDI imaging mass spectrometry of N-linked glycans on formalin-fixed paraffin-embedded murine kidney*. Anal Bioanal Chem, 2015. **407**(8): p. 2127-39.
266. Toghi Eshghi, S., et al., *Imaging of N-linked glycans from formalin-fixed paraffin-embedded tissue sections using MALDI mass spectrometry*. ACS Chem Biol, 2014. **9**(9): p. 2149-56.
267. Powers, T.W., et al., *Matrix assisted laser desorption ionization imaging mass spectrometry workflow for spatial profiling analysis of N-linked glycan expression in tissues*. Anal Chem, 2013. **85**(20): p. 9799-806.
268. 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*. Journal of Proteome Research, 2008. **7**(4): p. 1650-1659.
269. Norton, P., et al., *Development and application of a novel recombinant Aleuria aurantia lectin with enhanced core fucose binding for identification of glycoprotein biomarkers of hepatocellular carcinoma*. Proteomics, 2016. **16**(24): p. 3126-3136.
270. Houser, J., et al., *Influence of Trp flipping on carbohydrate binding in lectins. An example on Aleuria aurantia lectin AAL*. PLoS One, 2017. **12**(12): p. e0189375.
271. Blomme, B., et al., *Alteration of protein glycosylation in liver diseases*. J Hepatol, 2009. **50**(3): p. 592-603.
272. Yin, H., et al., *Mass-Selected Site-Specific Core-Fucosylation of Serum Proteins in Hepatocellular Carcinoma*. J Proteome Res, 2015. **14**(11): p. 4876-84.
273. Thakolwiboon, S., et al., *Heterogeneity of The CD90+ Population in Different Stages of Hepatocarcinogenesis*. J Proteomics Bioinform, 2014. **7**(10): p. 296-302.
274. Zhu, J., et al., *Analysis of serum haptoglobin fucosylation in hepatocellular carcinoma and liver cirrhosis of different etiologies*. J Proteome Res, 2014. **13**(6): p. 2986-97.
275. Norton, P., Comunale, M., Herrera, H., Wang, M., Houser, J., Wimmerova, M., Romano, P.R., Mehta, A., *Development and application of a novel recombinant Aleuria aurantia lectin with enhanced core fucose binding for identification of glycoprotein biomarkers of hepatocellular carcinoma*. Proteomics, 2016. **16**(24).
276. Romano, P.R., Mackay A., Vong, M. et al., *Development of recombinant Aleuria aurantia lectins with altered binding specificities to fucosylated glycans*. Biochemical and biophysical research communications, 2011. **414**(1): p. 84-89.
277. Block, T.M., et al., *Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans*. Proc Natl Acad Sci U S A, 2005. **102**(3): p. 779-84.

278. Johnson, P.J., et al., *Structures of disease-specific serum alpha-fetoprotein isoforms*. Br J Cancer, 2000. **83**(10): p. 1330-7.
279. Johnson, P.J., et al., *Glycan composition of serum alpha-fetoprotein in patients with hepatocellular carcinoma and non-seminomatous germ cell tumour*. Br J Cancer, 1999. **81**(7): p. 1188-95.
280. Ajdukiewicz, A.B., et al., *Alpha-fetoprotein glycosylation is abnormal in some hepatocellular carcinoma, including white patients with a normal alpha-fetoprotein concentration*. Cancer Lett, 1993. **74**(1-2): p. 43-50.
281. Murakami, I., et al., *Lectin binding patterns in normal liver, chronic active hepatitis, liver cirrhosis, and hepatocellular carcinoma. An immunohistochemical and immunoelectron microscopic study*. Acta Pathol Jpn, 1992. **42**(8): p. 566-72.
282. Furukawa, J., et al., *Comprehensive Glycomics of a Multistep Human Brain Tumor Model Reveals Specific Glycosylation Patterns Related to Malignancy*. PLoS One, 2015. **10**(7): p. e0128300.
283. Elder, J.H. and S. Alexander, *endo-beta-N-acetylglucosaminidase F: endoglycosidase from Flavobacterium meningosepticum that cleaves both high-mannose and complex glycoproteins*. Proceedings of the National Academy of Sciences of the United States of America, 1982. **79**(15): p. 4540-4544.
284. Trimble, R.B. and A.L. Tarentino, *Identification of distinct endoglycosidase (endo) activities in Flavobacterium meningosepticum: endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans*. The Journal of biological chemistry, 1991. **266**(3): p. 1646-1651.
285. Tarentino, A.L., et al., *Multiple endoglycosidase F activities expressed by Flavobacterium meningosepticum endoglycosidases F2 and F3. Molecular cloning, primary sequence, and enzyme expression*. The Journal of biological chemistry, 1993. **268**(13): p. 9702-9708.
286. Plummer, T.H., Jr., A.W. Phelan, and A.L. Tarentino, *Porcine fibrinogen glycopeptides: substrates for detecting endo-beta-N-acetylglucosaminidases F2 and F3(1)*. Analytical biochemistry, 1996. **235**(1): p. 98-101.
287. Norton, P., et al., *Development and application of a novel recombinant *Aleuria aurantia* lectin with enhanced core fucose binding for identification of glycoprotein biomarkers of hepatocellular carcinoma*. PROTEOMICS, 2016. **16**(24): p. 3126-3136.
288. Campbell, M.P., et al., *GlycoBase and autoGU: tools for HPLC-based glycan analysis*. Bioinformatics, 2008. **24**(9): p. 1214-1216.
289. Powers, T.W., et al., *MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays*. PLoS One, 2014. **9**(9): p. e106255-e106255.
290. Angel, P.M., et al., *MALDI Imaging Mass Spectrometry of N-glycans and Tryptic Peptides from the Same Formalin-Fixed, Paraffin-Embedded Tissue Section*. Methods Mol Biol, 2018. **1788**: p. 225-241.

291. Karamessinis, P.M., et al., *Marked Defects in the Expression and Glycosylation of 2-HS Glycoprotein/Fetuin-A in Plasma from Neonates with Intrauterine Growth Restriction* *PROTEOMICS SCREENING AND POTENTIAL CLINICAL IMPLICATIONS\**. *Molecular & Cellular Proteomics*, 2008. **7**: p. 591-599.
292. Lin, Y.-H., V. Franc, and A.J.R. Heck, *Similar Albeit Not the Same: In-Depth Analysis of Proteoforms of Human Serum, Bovine Serum, and Recombinant Human Fetuin*. *Journal of Proteome Research*, 2018. **17**(8): p. 2861-2869.
293. Prien, J.M., et al., *The high mannose glycans from bovine ribonuclease B isomer characterization by ion trap MS*. *J Am Soc Mass Spectrom*, 2009. **20**(4): p. 539-56.
294. Miyoshi, E., et al., *The alpha1-6-fucosyltransferase gene and its biological significance*. *Biochimica et biophysica acta*, 1999. **1473**(1): p. 9-20.
295. Mehta, A., H. Herrera, and T. Block, *Glycosylation and liver cancer*. *Advances in cancer research*, 2015. **126**: p. 257-79.
296. Wongtrakul-Kish, K., et al., *Combining Glucose Units, m/z and Collision Cross Section values: Multi-attribute data for increased accuracy in automated glycosphingolipid glycan identifications and its application in Triple Negative Breast Cancer*. *Analytical Chemistry*, 2019.
297. Nishikaze, T., *Sialic acid derivatization for glycan analysis by mass spectrometry*. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, 2019. **95**(9): p. 523-537.
298. Pearce, O.M.T. and H. Läubli, *Sialic acids in cancer biology and immunity*. *Glycobiology*, 2015. **26**(2): p. 111-128.
299. Büll, C., et al., *Sialic Acid Blockade Suppresses Tumor Growth by Enhancing T-cell-Mediated Tumor Immunity*. *Cancer Research*, 2018. **78**(13): p. 3574.
300. Vajaria, B.N., et al., *Sialylation: an Avenue to Target Cancer Cells*. *Pathology & Oncology Research*, 2016. **22**(3): p. 443-447.
301. Schauer, R. and J.P. Kamerling, *Chapter One - Exploration of the Sialic Acid World*, in *Advances in Carbohydrate Chemistry and Biochemistry*, D.C. Baker, Editor. 2018, Academic Press. p. 1-213.
302. Lu, J. and J. Gu, *Significance of  $\beta$ -Galactoside  $\alpha$ 2,6 Sialyltransferase 1 in Cancers*. *Molecules (Basel, Switzerland)*, 2015. **20**(5): p. 7509-7527.
303. Pihikova, D., et al., *Sweet characterisation of prostate specific antigen using electrochemical lectin-based immunosensor assay and MALDI TOF/TOF analysis: Focus on sialic acid*. *Proteomics*, 2016. **16**(24): p. 3085-3095.
304. Dédová, T., et al., *Sialic Acid Linkage Analysis Refines the Diagnosis of Ovarian Cancer*. *Frontiers in oncology*, 2019. **9**: p. 261-261.
305. Dall'Olio, F., et al., *Expression of  $\beta$ -galactoside  $\alpha$ 2,6 sialyltransferase and of  $\alpha$ 2,6-sialylated glycoconjugates in normal human liver, hepatocarcinoma, and cirrhosis*. *Glycobiology*, 2004. **14**(1): p. 39-49.

306. Hein, C.D., X.-M. Liu, and D. Wang, *Click chemistry, a powerful tool for pharmaceutical sciences*. *Pharmaceutical research*, 2008. **25**(10): p. 2216-2230.
307. Chen, R., et al., *Site-specific characterization of cell membrane N-glycosylation with integrated hydrophilic interaction chromatography solid phase extraction and LC-MS/MS*. *Journal of Proteomics*, 2014.
308. Chen, Z., et al., *Site-specific characterization and quantitation of N-glycopeptides in PKM2 knockout breast cancer cells using DiLeu isobaric tags enabled by electron-transfer/higher-energy collision dissociation (EThcD)*. *Analyst*, 2018. **143**(11): p. 2508-2519.
309. Levery, S.B., et al., *Advances in mass spectrometry driven O-glycoproteomics*. *Biochimica et Biophysica Acta - General Subjects*, 2015.
310. Liu, M.-Q., et al., *pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification*. *Nature Communications*, 2017. **8**(1): p. 438.
311. Renee Ruhaak, L., et al., *Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses*. *Chemical Review*.
312. Thaysen-Andersen, M. and N.H. Packer, *Advances in LC-MS/MS-based glycoproteomics: Getting closer to system-wide site-specific mapping of the N- and O-glycoproteome*. 2014.
313. Yin, H., et al., *Mass-selected site-specific core-fucosylation of ceruloplasmin in alcohol-related hepatocellular carcinoma*. *Journal of proteome research*, 2014. **13**(6): p. 2887-2896.
314. Zhang, Z., et al., *High-Throughput Determination of the Site-Specific N-Sialoglycan Occupancy Rates by Differential Oxidation of Glycoproteins Followed with Quantitative Glycoproteomics Analysis*.
315. Zhu, J., et al., *Glycoproteomic markers of hepatocellular carcinoma-mass spectrometry based approaches*. *Mass Spectrometry Reviews*, 2019. **38**(3): p. 265-290.
316. Zhou, S., et al., *Quantitative LC-MS/MS Glycomic Analysis of Biological Samples Using AminoxyTMT*. *Analytical chemistry*, 2016. **88**(15): p. 7515-22.
317. Alley Jr, W.R., Y. Mechref, and M.V. Novotny, *Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data*. *Rapid Communications in Mass Spectrometry*, 2009. **23**(1): p. 161-170.
318. Ruiz May, E., C. Catalá, and J. Rose, *N-Glycoprotein Enrichment by Lectin Affinity Chromatography*. *Methods in molecular biology (Clifton, N.J.)*, 2014. **1072**: p. 633-43.
319. Sprung, R.W., et al., *Precision of multiple reaction monitoring mass spectrometry analysis of formalin-fixed, paraffin-embedded tissue*. *Journal of proteome research*, 2012. **11**(6): p. 3498-3505.
320. Benicky, J., et al., *Quantification of fucosylated hemopexin and complement factor H in plasma of patients with liver disease*. *Analytical chemistry*, 2014. **86**(21): p. 10716-10723.



321. Chauhan, R. and N. Lahiri, *Tissue-and Serum-Associated Biomarkers of Hepatocellular Carcinoma Supplementary Issue: Biomarkers and their Essential Role in the Development of Personalised Therapies*. Biomarkers in CanCer, 2016. **8**(s1).
322. Goldman, R., et al., *Detection of hepatocellular carcinoma using glycomic analysis*. Clinical cancer research : an official journal of the American Association for Cancer Research, 2009. **15**(5): p. 1808-13.
323. Mehta, A. and T.M. Block, *Fucosylated glycoproteins as markers of liver disease*. 2008.
324. Gao, H.-J., et al., *Quantitative proteomic analysis for high-throughput screening of differential glycoproteins in hepatocellular carcinoma serum*. Cancer biology & medicine, 2015. **12**(3): p. 246-254.
325. Attallah, A.M., et al., *Diagnostic role of collagen-III and matrix metalloproteinase-1 for early detection of hepatocellular carcinoma*. British Journal of Biomedical Science, 2020. **77**(2): p. 58-63.
326. Cheung, P. and J.W. Dennis, *Mgat5 and Pten interact to regulate cell growth and polarity*. 2007. **17**(7): p. 767-773.