2015

Developing an Integrative Glycobiology Workflow for the Identification of Disease Markers for Pancreatic Cancer

Thomas Wesley Powers
Medical University of South Carolina

Developing an Integrative Glycobiology Workflow for the Identification of Disease Markers for Pancreatic Cancer

by

Thomas Wesley Powers

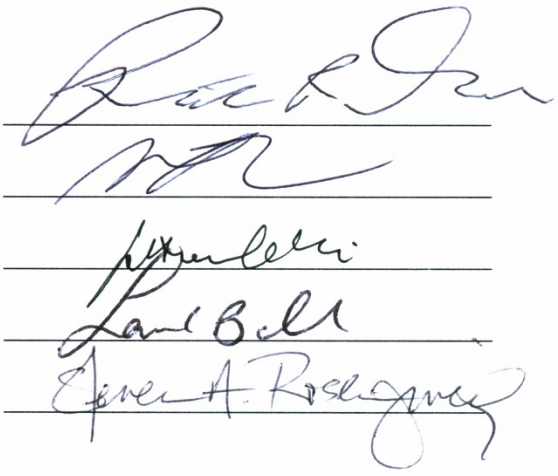
A dissertation submitted to the faculty of the Medical University of South Carolina in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the College of Graduate Studies

Department of Cell and Molecular Pharmacology and Experimental Therapeutics

2015

Approved by:

Chairman, Advisory Committee



ABSTRACT

THOMAS WESLEY POWERS. Developing and Integrative Glycobiology Workflow for the Identification of Disease Markers for Pancreatic Cancer (Under the Direction of Richard Drake)

A deeper understanding of dysregulated glycosylation in pancreatic cancer can provide insights into disease mechanisms and the identification of novel disease markers. Recent improvements in mass spectrometry techniques have been instrumental in profiling biologically relevant tissue sections in order to identify disease marker candidates, but have either not yet been adopted for studying glycosylation or applied directly to pancreatic cancer. In the dissertation herein, new methods have been developed and adapted to the study of aberrant glycosylation in pancreatic cancer, with the ultimate goal of identifying novel disease marker candidates. For the first time, we describe a mass spectrometry imaging approach to study the localization of N-glycans. This technique demonstrated a histology-derived localization of N-glycans across tissue sections, with identifications displaying remarkable consistency with documented studies. Furthermore, the technique provides superior structural information compared to preexisting methodologies. In the analysis of diseased specimen, changes in glycosylation can be linked to aberrations in glycosyltransferase expression. When applied to pancreatic cancer in a high-throughput and high-dimensional analysis, panels of glycans displayed an improved ability to differentiate tumor from non-tumor tissues compared to current disease markers. Furthermore, the data suggest that glycosylation can identify premalignant lesions, as well as differentiate between malignant and benign conditions. These observations overcome significant limitations that hinder the efficacy of current disease markers. In an effort to link aberrant glycosylation to the modified protein, a subset of glycosylated proteins were enriched and analyzed by mass spectrometry to identify proteins that are integral to disease progression and can be probed for the early detection

of pancreatic cancer. Known disease markers were among the glycoproteins identified, validating the utility of the enrichment and detection strategy outlined. This approach also differentiated the role of N- and O-glycosylation in antigen expression. Finally, we outline an integrated workflow that takes advantage of the unique capabilities of high resolution mass spectrometers. This workflow can capitalize on prior glycomic and proteomic experiments to provide a comprehensive analysis of dysregulated protein glycosylation in pancreatic cancer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to a number of people, all of whom have provided me the support and encouragement necessary to complete this research. I would like to thank my mentor, Dr. Richard Drake, for the constant guidance, support, and intellectual input related to my project. Additionally, I am grateful for the opportunities to participate in collaborative research projects and attend national conferences, which have allowed me to grow as a scientist and independent thinker. His personality and kindness have been an inspiration to me and made working in his lab a great experience. A special thanks to the rest of the members of my committee; Drs. Ball, Gattoni Celli, Kindy, and Rosenzweig. Each member assisted offered unique insights that allowed me to advance the scope of the project and overcome limitations. I appreciate the support of Drs. Haab, Lewin, Mehta, and Shao, for their contributions to the dissertation research. A special thanks to all members of the Drake lab, past and present, as well as members of the Ball lab and the mass spectrometry core facility.

I am personally indebted to numerous individuals. To my grandparents, who have always been a source of inspiration to me, both professionally and personally. To my sister, who provided me with a lofty benchmark of success for which to strive. I am eternally grateful to my parents for their constant love, support, encouragement and belief that they have shown in me. I am extremely priveledged to have such a loving and compassionate family. Finally, I would like to thank my wife and daughter. I would like to thank Virginia for her patience and understanding. While it was not always easy, especially the last few hectic months, her fortitude during the process has been greatly appreciated. Her belief in me, especially in times of personal doubt, have constantly motivated and inspired me to keep going. Finally, my sincerest appreciation to my daughter Avery, who has been a source of immense joy during the process of writing this dissertation. To Virginia and Avery, this dissertation and the research herein is written in honor of you.

Chapter 1: Literature Review

1.1 Introduction to Pancreatic Cancer

1.1.1 Symptoms, Mortality, and Standard of Care

As annual cancer death rates continue to decline for a majority of cancers due to a reduction in disease-causing behaviors, improvements in early detection and screening approaches, and improved treatments for the major cancer sites, pancreatic cancer has seen a rise in death rates (1,2). Average 5-year survival rates for pancreatic cancer remain at a dismal level, around 6%, and the disease accounts for the fourth most cancer related deaths in the United States (2). For pancreatic cancer, poor clinical outcome is related to a poor understanding of the disease, rapid progression, limited effective treatment options, and resistance to therapeutic interventions. These factors are further enhanced by a pattern of delayed diagnosis, attributed to mild and ambiguous clinical presentations. As the pancreas is located deep in the abdomen, tumor growth may persist for extended periods of time before pressure or discomfort is detected. In addition to pain and discomfort, other symptoms include jaundice, digestive difficulties, ascites, unexpected weight loss and sudden onset diabetes (3,4). However, these symptoms are not specific to pancreatic cancer, and are therefore often attributed to other diseases or conditions. For these reasons, many patients remain asymptomatic until the disease has already progressed to late stages. The difficulty in detecting pancreatic cancer translates to approximately 9% of patients being diagnosed with localized cancer (5). For the remaining patients, the cancer has already spread to regional lymph nodes or metastasized to distant sites by the time of detection.

Unfortunately, 5-year survival rates and median lengths of survival are dramatically impacted by the stage of pancreatic cancer at the time of diagnosis. Stages of pancreatic cancer range from stage 0 to stage IV, with certain stages being further divided (6). The stage is determined at the time of diagnosis and is primarily based on the size of the tumor and the spread of the cancer to other sites (Table 1). Stage 0, or carcinoma in situ, refers to abnormal cells found

in the pancreas that may become cancerous. Pancreatic cancer is classified as stage IA if the tumor is less than 2 centimeters and is not found outside of the pancreas. Stage IV pancreatic cancer refers to cancer that has metastasized to distant organs (6). For the largest subtype of pancreatic cancer, 5-year survival decreases from 14% when the patient is diagnosed at with stage IA pancreatic cancer, to 1% when the patient is diagnosed with stage IV cancer (7). The greatest benefit of detecting pancreatic cancer early is the potential for complete surgical resection of the tumor, which is the only chance of curative treatment. However, surgical resection is only available for the 15-20% of pancreatic cancer patients that are deemed able to survive the surgery and are determined not to have locally invasive or metastatic cancer (3,8). Even with surgical resection, 5-year survival rates remain below 25% (2). A majority of tumors of the pancreas head and proximal neck are resected using the Whipple procedure followed by adjuvant chemotherapy (3,8). Patients with more advanced cancer are limited to chemotherapy. Gemcitabine, alone or in combination with 5-fluorouracil (5-FU), has traditionally been considered the standard of care chemotherapy for patients with locally advanced or metastatic cancer (8). However, gemcitabine treatment results in only modest patient benefits due to limited drug uptake and acquired resistance (9). Newer chemotherapy combinations, FOLFIRINOX and a combination of gemcitabine plus nanoparticle albumin-bound paclitaxel (nab-paclitaxel), have been adopted as the new standard of care based off of improved survival statistics (8). FOLFIRINOX, a combination of leucovorin, 5-FU, irinotecan and oxaliplatin, showed an increase in survival to 11.1 months compared to 6.8 months for gemcitabine alone (10). Unfortunately, the toxicity of FOLFIRINOX limits the therapeutic use. The combination of gemcitabine and nab-paclitaxel outperformed gemcitabine alone in survival and progression-free survival (11). While therapeutic intervention has been improved for more advanced stages of pancreatic cancer, these efforts have only a modest effect on life expectancy.

Table 1. Stage, Resectability and Survival of Pancreatic Cancer

	Clinical Stage	Resectable	Median Survival	Diagnosis	5-Year Survival
Local	Stage 0 (Tis, N0, M0)	Yes	17-23 Mos	10%	13%
	Stage IA (T1, N0, M0)				
	Stage IB (T2, N0, M0)				
	Stage IIA (T3, N0, M0)				6%
	Stage IIB (T1-3, N1, M0)				
Locally Advanced	Stage III	Borderline	<20 Mos	30%	3%
	Stage III (T4, any N, M0)	No	8-14 Mos		
Metastatic	Stage IV (any T, any N, M1)	No	4-6 Mos	60%	1%

Resection: 18-24%

T = Primary tumor
Tis: Carcinoma in situ
T1: Tumor restricted to pancreas, <2cm
T2: Tumor restricted to pancreas, >2 cm
T3: Tumor extends beyond pancreas
T4: Tumor affects coeliac axis or superior mesenteric artery

N = Regional lymph node
N0: No regional lymph-node metastasis
N1: Regional lymph-node metastasis

M = Distant metastasis
M0: No distant metastasis
M1: Distant metastasis

1.1.2 Pancreatic Cancer Subtypes

The pancreas is characterized as having both an exocrine and endocrine function. Exocrine cells, which constitute a majority of the cells in the pancreas, are responsible for the production and release of pancreatic juices involved in digestion. A majority of pancreatic cancers are exocrine in nature; including ductal adenocarcinomas, cystic tumors and cancer of acinar cells. Among the forms of exocrine pancreatic cancer, pancreatic ductal adenocarcinoma (PDAC) is estimated to account for 90% of all pancreatic cancer instances. Ultimately, PDAC is the exocrine tumor with the worst prognosis, as other exocrine tumors are slower growing and have higher 5-year survival rates (12). The second most common subtype, cystic tumors, account for less than 5% of all pancreatic cancers (13). While many of these cystic lesions are benign, lesions with malignant potential include mucinous cystic neoplasms (MCN), serous cystadenomas, and intraductal papillary mucinous tumors (IPMT). MCNs are characterized by excessive mucin production while IPMNs resemble early stage PDAC lesions but form larger cystic structures (14,15). Apart from morphological differences, the location of the lesion within the pancreas can aid in diagnosis (12). For patients with exocrine cancer, both the cancer and

surgical intervention may result in exocrine insufficiency, a condition marked by a deficiency in digestive enzymes produced in the exocrine pancreas (16,17). Treatment of these patients with pancreatic enzyme replacement therapy can offset unintentional weight loss caused by maldigestion and alleviate some symptoms of the disease (18).

Pancreatic cancer of the endocrine pancreas, whose biological function is to modulate glucose regulation, occurs infrequently and constitutes less than 5% of the instances of pancreatic cancer. Endocrine tumors, often referred to as neuroendocrine or islet cell tumors, are most often indolent but have malignant potential if left untreated. Endocrine tumor metastasis to the liver is the most common site of metastasis and the primary cause of death for these patients (19). Endocrine tumors are classified as functional, where the tumor produces excessive amounts of hormones; or nonfunctional, where the tumor does not secrete excess hormones. Individuals with functional neuroendocrine tumors present with symptoms reflective of the hypersecreted hormone. Functional tumors include insulinoma and glucagonomas, which secrete excessive amounts of insulin and glucagon, respectively. Unlike functional tumors, which present with symptoms specific to the hypersecreted hormone, nonfunctional tumors are less symptomatic and are frequently discovered incidentally or at more advanced stages (20). The distinction between functional and nonfunctional tumors is based on symptoms at the time of diagnosis as opposed to concrete hormonal expression differences, meaning that nonfunctional tumors can eventually trigger symptoms associated with functional endocrine tumors. Recently, treatment strategies for endocrine tumors have transformed from a "wait-and-see" approach to a more aggressive treatment approach (19). While still poor, 5-year survival is markedly higher for individuals with endocrine tumors compared to exocrine pancreatic tumors. The 5-year survival rates for exocrine tumors, largely biased by the high relative occurrence and poor outcome of PDAC, are 14%, 12%, 7%, 5%, 3%, and 1% for stage IA, stage IB, stage IIA, stage IIB, stage III, and stage IV, respectively (7). In contrast, 5-year survival rates for patients with neuroendocrine tumors

following surgery are 61%, 52%, 41%, and 16% for stage I, stage II, stage III, and stage IV tumors, respectively (7). The 5-year survival of individuals with exocrine tumors that are eligible for surgical resection remains markedly lower than that of patients with endocrine tumors, except those diagnosed with the worst stage of endocrine cancer.

1.1.3 Disease Etiology

1.1.3.1 Risk Factors

Both environmental cues and genetic predispositions are associated with the development of pancreatic cancer. From a population standpoint, pancreatic cancer is most often diagnosed in individuals between 60-80 years of age, and rarely before the age of 40 (12). Men are more likely to develop pancreatic cancer, with rates for Caucasians being lower than other races (21). After controlling for enviromentental factors, studies have established in increased risk of developing pancreatic cancer associated with smoking, chronic pancreatitis, excessive alcohol consumption, type-2 diabetes, and obesity (22–28). These factors can induce pancreatic cancer through different mechanisms. Pancreatitis results in the release of growth factors, cytokines, and reactive oxygen species, while smoking, diabetes and obesity result in chronic inflammatory states (14,26). These environmental effects are often interconnected, such as heavy alcohol consumption promoting chronic pancreatitis (27,28). Among the controllable risk factors, smoking has the most profound effect on the population, and is estimated to result in a 2-fold increase in the rate of pancreatic cancer development and accounts for approximately 25% of all instances of PDAC (21). Efforts to identify the impact of dietary factors on PDAC incidence rates have been largely contested, with the most likely impact being through obesity and diabetes (21). However, there is some debate as to whether diabetes is in fact a cause of pancreatic cancer. Patients with new-onset type-2 diabetes, as opposed to long time type-2 diabetic patients, have a higher risk of developing PDAC, suggesting the onset of diabetes may instead be a manifestation of PDAC (29).

Genetic predispositions are estimated to cause around 10% of the instances pancreatic cancer (30,31). Individuals with 1, 2, and 3 first degree relatives with pancreatic cancer have a greater than 3-, 4-, and 14-fold increased risk of developing pancreatic cancer compared to individuals with no familial history of pancreatic cancer (26). Risk is further elevated when there is a history of early onset pancreatic cancer (age < 50 years) (32). *BRCA1*, *BRCA2*, *PALB2*, *STK11*, *CDKN2A, PRSS1,* and *TP53* gene mutations are associated with elevated instances of PDAC (33–36). Among these, individuals with *STK11* and *CDKN2A* mutations have the highest rate increase for developing PDAC compared to individuals without the gene mutations. Germline *BRCA2* gene mutations contribute to the most instances of hereditary pancreatic cancer (32). Many of the genes mentioned above are frequently mutated in pancreatic cancer, and will be discussed further in regards to their role in cell signaling. While chronic pancreatitis is often induced by excessive alcohol consumption and thus characterized as an environmental risk factor, hereditary pancreatitis is associated with a 40% an increased risk of acquiring pancreatic cancer (21). It is likely that there are additional unknown genetic predispositions for pancreatic cancer.

1.1.3.2 Common Signaling Pathways Linked to Pancreatic Cancer

While the exact biological mechanisms that drive of pancreatic cancer initiation and progression are not fully understood, recent studies have identified consistent molecular pathways that are likely involved. These mutations appear to regulate different stages of malignant progression, some appearing in early stage pancreatic intraepithelial neoplasia (PanIN) lesions while others manifest only in later stage PanIN lesions and malignancy (Figure 1). While malignant progression typically proceeds though PanIN lesions, other initiating lesions exist in the form of IPMNs and MCNs (15). Activating *KRAS* mutations are observed in many cancers, including pancreatic cancer (37). Gain-of-function *KRAS* mutations on chromosome p12 are estimated to occur in over 90% of pancreatic cancers and 36%, 44% and 87% of PanIN1, PanIN2, and PanIN3 lesions, respectively (15,38). While the substitution mutation of glycine for aspartate

at codon 12 is the most common mutation, other single point mutations at residues G12, G13, and Q61 have been observed (39,40). Mechanistically, substitutions at G12 and G13 produce steric interference between KRAS and GTPase activating proteins (GAPs), prolonging KRAS activation and enhancing downstream signaling (40,41). Under normal conditions, KRAS localizes to the plasma membrane and mediates a variety of biological processes, including proliferation, differentiation and survival (14,37). The prolonged KRAS activation in pancreatic cancer results in increased proliferation, suppression of apoptosis, alterations in the metabolism of the tumor microenvironment, and metastasis (40). While the exact KRAS-mediated pathways utilized during pancreatic cancer tumorigenesis are unknown, the mitogen-activated protein kinase (MAPK), phosphoinositide 3-kinase (PI3K) and ral guanine nucleotide exchange factor (RalGEFs) pathways have all been implicated. KRAS-mediated MAPK/ERK activation results in a proliferative, anti-apoptotic and invasive cellular phenotype in PDAC cell lines (14,42). Furthermore, in a murine model, the selective activation of the Raf-MEK-ERK pathway by oncogenic $Braf^{V600E}$ induced a more aggressive pancreatic cancer phenotype with more PanIN lesions than the $KRAS^{G12D}$ murine model (43). Similarly, KRAS-mediated PI3K/AKT signaling regulates cell growth, motility and survival. PDAC is associated with increased signaling attributed to KRAS activity, decreased expression of PTEN, or amplification of AKT2 (14,44,45). In murine models, expression of the oncogenic $PIK3CA^{H1047R}$ allele, which activates expression of $p110\alpha^{H1047R}$ and PI3K signaling in the pancreas, resulted in PDAC tumors with nearly identical survival and metastasis rates as $KRAS^{G12D}$ mutants (46). Additionally, blocking downstream PI3K signaling by inactivating *PDK1* stopped the formations of PanIN and PDAC in $KRAS^{G12D}$ mouse models(46). From a therapeutic standpoint, targeting downstream intermediates of these pathways have been more fruitful than KRAS inhibitors. Inhibition of downstream effectors of the MAPK or PI3K pathways, including MEKK, AKT2 and mTOR, results in growth inhibition, increased sensitivity to chemotherapy and reduced invasiveness in PDAC cell lines

and xenograft assays (26,45,47,48). While *KRAS* mutations appear early in tumorigenesis, they are frequently absent in PanIN lesions that ultimately progress to pancreatic cancer, and present in non-cancerous PanIN lesions, suggest that other signaling pathways are involved in the induction of pancreatic cancer, likely working in conjunction with KRAS signaling.

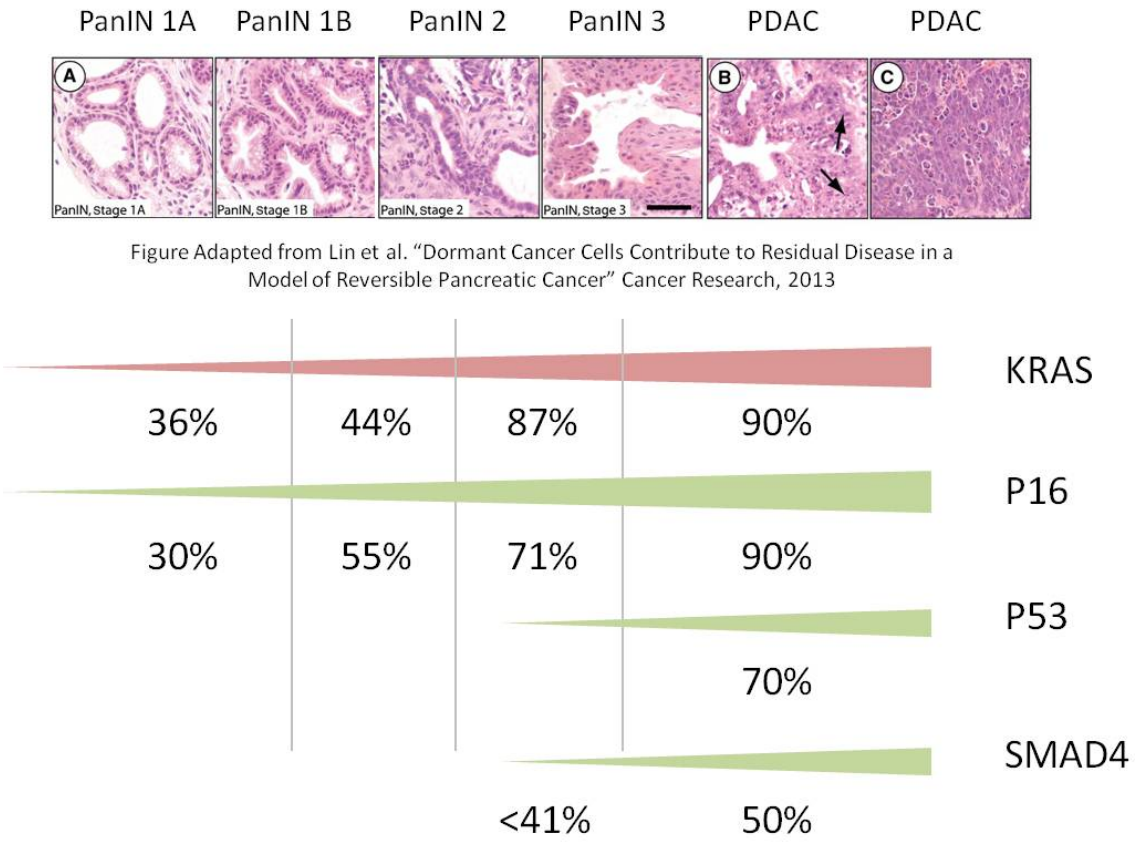


Figure 1. PanIN to PDAC Transition is Associated with Genetic Mutations. The transition of the cellular morphology that occurs from the progression early PanIN lesions to full blown pancreatic cancer is accompanied with mutations of the KRAS oncogene and the P16, P53 and SMAD4 tumor suppressors.

In addition to *KRAS* oncogene mutations, three tumor suppressors are frequently inactivated in pancreatic cancer; these being *CDKN2A, TP53* and *SMAD4*. Inactivation of the p16/INK4/CDKN2A tumor suppressor in pancreatic cancer occurs through homozygous deletion,

intragenic mutation, and hypermethylation of the promoter in 40%, 40% and 15% of instances, respectively (49). The resulting loss of p16 function occurs in around 90% of pancreatic cancers and is also observed in 30% of PanIN1 lesions, 55% of PanIN2 lesions, and 71% of PanIN3 lesions (15,38). While this mutation occurs sporadically in pancreatic cancer, germline mutations of the *p16/INK4/CDKN2A* gene is associated with familial atypical mole-malignant melanoma (FAMMM) syndrome, a condition that is linked to a 13-fold increased risk of pancreatic cancer (14). Traditionally, the tumor suppressor gene *p16/INK4/CDKN2A* regulates cell cycle progression during G1/S transition of the cell cycle. Cancer derived loss of p16 function increases Rb phosphorylation via the interaction of CDK4/6 with Cyclin D1, resulting in the release of E2F transcription factors and gene transcription required for S Phase (50). However, an additional role of inactivation of the *p16/INK4/CDKN2A* gene has been hypothesized. It is now known that $p14^{ARF}$ is located on the same gene locus, which functions to regulate p53-mediated cell cycle arrest through the stabilization of p53 (14,15,50). As the *p16/INK4/CDKN2A* locus encodes for two tumor suppressors, p16 and $p14^{ARF}$, efforts to identify which gene plays a role in pancreatic tumorigenesis are ongoing. Inactivating mutations that target *p16* or *p16* and $p14^{ARF}$, emphasize the importance of p16 as a tumor suppressor (33). Meanwhile, an analysis of pancreatic cancer mouse models established a role for both *p16* and $p19^{ARF}$, the equivalent gene to $p14^{ARF}$ in humans, in pancreatic cancer progression (38). Mutations in p16 may also be responsible for avoiding senescence induced by *KRAS* oncogene signaling (14,15,46).

P53 is mutated in pancreatic cancer, diminishing the ability of the transcription factor to regulate the G1/S and G2/M phases of cell-cycle progression and induce apoptosis. Uniquely, pancreatic cancer commonly diminishes p53 function by two mechanisms; $p14^{ARF}$ deficiency as a result of certain *CDKN2A* mutations or missense mutations in *TP53* that encodes for the DNA-binding domain (15,51). In the first mechanism, mutations of the $p14^{ARF}$ gene elevates expression of MDM2 and proteolysis of p53 (14). Alternatively, mutations in p53 are estimated to occur in

over 70% of pancreatic cancers and in late stage PanIN lesions. The mutated protein forms heterooligomers with wild-type p53, resulting in impaired association with DNA and transcriptional activation (52–54). Unlike many cancers where these mechanisms are mutually exclusive, *TP53* mutations and $p14^{ARF}$ deficiencies are both present in around 40% of pancreatic cancer patients (14). However, there are additional oncogenic, 'gain-of-function' properties that are observed only in the presence of mutated p53, suggesting additional functions for mutated p53 in pancreatic cancer progression. These properties result in chromosomal instability, the inhibition of DNA repair pathways, and more aggressive and drug resistant cell phenotypes (55). This aggressive phenotype is demonstrated by the observation that 65% of mice expressing mutated p53 developed metastases, while none of the p53-null mice did (56).

DPC4/SMAD4/MADH4 is the third tumor suppressor that is commonly mutated in pancreatic cancer. Like p53 mutations, loss of SMAD4 is seen late in pancreatic cancer progression, appearing in 31-41% of PanIN3 lesions and 50% of cancers (14,15). The loss of SMAD4 is thought to contribute to pancreatic cancer through TGFβ signaling. Under normal conditions, TGFβ binds cell surface receptors and initiates an intracellular kinase cascade resulting in the phosphorylation and translocation of SMAD4 to the nucleus, where SMAD4 transcriptionally regulates apoptosis, proliferation and migration. While TGFβ exerts growth-inhibitory and growth-promoting effects, it is thought that the loss of SMAD4 results in decreased growth inhibition in pancreatic cancer. Additionally, the loss of SMAD4 does not reduce the TGFβ-mediated promotion of EMT, shifting TGFβ to a tumor promoter (52,57). This is in accordance with the *in vivo* observation of reduced invasion upon TGFβ inhibition, elevated invasion and EMT upon addition of exogenous TGFβ, and elevated levels of TGFβ in PDAC cancer cells compared to normal cells and the effect (58–60).

Other signaling pathways and biological events have been also been implicated in pancreatic cancer, albeit not as frequently as those described above. The tumor suppressors

LKB1/STK11 and *BRCA2* can be mutated in both sporadic and familial pancreatic cancer (61,62). Peutz-Jeghers syndrome, characterized by *LKB1/STK11* gene mutation, is associated with a >40%-fold increase in PDAC instances (14). Whereas LKB1 promotes tumorigenesis through the regulation of mTOR, *BRCA2* mutations impair DNA damage response pathways (63,64). *BRCA2* gene mutations are observed in around 17% of pancreatic cancers, and present in later staged lesions, suggesting a role in malignant transformation of precancerous lesions to full-blown PDAC (14,62). Other studies have implicated constitutive NF-kB signaling in pancreatic cancer based on the observation that inhibition of NF-kB inhibits tumorigenesis and increases sensitivity to anticancer agents (65–67). Similarly, both notch and sonic hedgehog developmental signaling pathways, which are absent in the normal adult pancreas, are activated in PanIN lesions and PDAC (68,69). Finally, both growth factors (EGF, IGF, VEGF) and their corresponding receptors have also been implicated in pancreatic cancer growth and proliferation (70–74). The disruption of growth factor receptor signaling has been shown to suppress both in vivo and in vitro tumor growth and enhance response to chemotherapy (74–78).

1.3.3.3 Origin of Pancreatic Cancer Initiation

The definitive origin of the cells responsible for pancreatic cancer initiation remains unclear. Given the ductal-like morphology of the cancer, ductal cells were originally hypothesized to be the cellular origin of PDAC (79). This theory is supported by evidence linking the transformation of neoplastic ductal lesions (PanINs) to pancreatic cancer, including similarities in genetic alterations. On the genetic level, approximately 30% of patients with PanIN-1A lesions exhibit point mutations in the *KRAS* oncogene, a mutation that becomes more prominent as the lesions ultimately progress to PDAC, where the mutation is present in about 90% of patients (80,81). Similarly, *CDKN2A*, *TP53*, and *SMAD4* exhibit inactivating mutations in PanIN lesions and, to a larger extent, pancreatic cancer (40). While evidence linking the early genetic mutations in PanIN with those frequently observed in PDAC supports the theory of ductal

cell origin, this concept is difficult to prove, especially as expression of oncogenic KRAS specifically in ductal cells does not induce PanIN or PDAC in murine models (82). On the other hand, more recent data suggests that acinar/centroacinar cells can acquire features of ductal cells and progress to PDAC (83,84). Transgenic murine models with ectopic expression of TGFα or c-myc demonstrate acinar cell acquisition of ductal cell features culminating in PDAC (84,85). Additionally, the insertion of activated KRAS into the open reading frame of *Mist*1, a transcription factor that is implicated in acinar organization, results in the development of aggressive pancreatic cancer in mice (86). Unlike ductal cells, the selective expression of oncogenic KRAS in acinar/centroacinar cells does induce PanINs and PDAC (83). Alternatively, subpopulations of cells that display characteristics analogous to stem cells have been hypothesized to act as cancer stem cells in pancreatic cancer. The subpopulation of pancreatic cancer cells that test positive for cell surface markers $CD44^+CD24^+ESA^+$ represent <1% of pancreatic cancer cells (87,88). These cells display highly tumorigenic properties and are estimated to have greater than 100-fold higher tumorgenicity compared to cells that test negative for the surface markers (88). In a xenograft mouse model, tumors that developed from the $CD44^+CD24^+ESA^+$ cells were indistinguishable from the parent tumor. Overall, the $CD44^+CD24^+ESA^+$ cells displayed properties of self-renewal, the potential to produce differentiated cells, and resulted in the expression of sonic hedgehog signaling molecules, all properties common in cancer stem cells (88). Alternatively, $CD133^+CXCR4^+$ cells influence the metastatic phenotype of pancreatic cancer cells (89). These cells are predominantly localized at the invasive front of the tumor and the different surface markers confer survival advantages to the cells. $CD133^+$ cells are more resistant to gemcitabine than $CD133^-$ cells, while the $CXCR4^+$ marker identifies migratory cells. Liver metastases were unique in mice given $CD133^+CXCR4^+$ cells compared to $CD133^+CXCR4^-$ cells (89).

1.1.4 Detection and Diagnosis of Pancreatic Cancer

1.1.4.1 Clinical Approaches to Detect Pancreatic Cancer

Pancreatic cancer is difficult to detect due to ambiguous symptoms, as described previously. When individuals present with symptoms, possess genetic mutations that increase their probability of developing pancreatic cancer, or have a family history of pancreatic cancer, a physician may perform a physical examination to check for jaundice, physical masses or fluid buildup in the abdomen. Non-invasive and invasive imaging tests provide more information regarding the status of pancreatic cancer. Abdominal ultrasounds are typically used for initial examinations but are often followed up by other non-invasive strategies, such as computed tomography scans (CAT), magnetic resonance imaging (MRI), magnetic resonance cholangiopancreatography (MRCP) or positron emission tomography (PET) scans (90). Depending on the approach, these imaging techniques can identify the presence of cancer, locate metastatic sites, and differentiate between benign and metastatic masses. Invasive imaging tests, termed endoscopic ultrasounds (EUS) or endoscopic retrograde cholangiopancreatography (ERCP) tests allow for more detailed images, but can also be used to insert stents to compensate for obstructed ducts or obtain fine needle aspirate (FNA) biopsy samples (12). EUS is capable of detecting small pre-invasive lesions, making it well suited to identify lesions in high-risk individuals (32). A combination of multiple screening strategies may be used to improve the accuracy of cancer detection (91). The combination of EUS, multi-detector computed tomography (MD-CT), MRI and MRCP offers the best sensitivity for the detection of pancreatic cancer (90). Absolute certainty in identification is only achievable with a biopsy. In addition to FNA biopsies, brush biopsies and tissue biopsies can be used for disease diagnosis. In a brush biopsy, cells are collected during an ERCP test, whereas tissue biopsies are collected during laparoscopy. In addition to imaging tests and biopsy collections, a blood test can be performed to

assess the probability of pancreatic cancer. Blood work can assess levels of bilirubin, pancreatic enzymes, CA19-9, and CEA. While altered levels of enzymes or markers in the blood work do not conclusively determine if an individual has pancreatic cancer, they can determine the necessity of further evaluations (92). After identification of pancreatic cancer, genetic testing may be used to identify gene mutations. While screening is useful in prompting immediate family members to undergo genetic testing, genetic testing is underutilized by physicians (32).

1.1.4.2 Gross and Microscopic Characteristics

Pancreatic cancer presents grossly as white-yellow firm masses, most often arising in the pancreatic head, but also infiltrating and metastasizing into surrounding tissues, including lymph nodes, spleen, peritoneal cavity, liver, and lungs (14,93). The tumor often has ill-defined edges and is surrounded by a dense fibrotic stroma and ECM proteins (79,94,95). The mean diameter of the tumor in the pancreatic head is between 2.5 and 3.5 cm, with larger tumors detected in the body or tail of the pancreas (12). At the microscopic level, the progression from low-grade (PanIN-1A or PanIN-1B) to high-grade (PanIN-3) lesions is accompanied by histological changes. Low-grade lesions are often flat and retain nuclear polarity, while high-grade lesions exhibit nuclear atypica, loss of nuclear polarity, and frequent mitoses (15). The transition from PanIN-3 lesions to invasive carcinoma is marked by the invasion through the basement membrane (93). The presence of cancer is frequently associated with infiltrating small glands lined with mucin-containing cells, with the cancer surrounded by a desmoplastic reaction (93). Malignant glands can be distinguished from benign glands based off on variations in the gland architecture, location of the glands, variation in the size of nuclei, association with surrounding pancreatic parenchyma, and the presence of necrotic debris (96). PDAC is typically moderately or well differentiated, but there is often some variation in the degree of differentiation within the tumor (12,97). As glands become less differentiated, the lumen boundary and overall shape of the

duct becomes more irregular (96). These glands commonly appear directly adjacent to muscular arteries or invade into perineural or vascular tissue.

1.2 Mass Spectrometry

1.2.1 Principles of Mass Spectrometry

1.2.1.1 General Introduction to Mass Spectrometry

Mass spectrometry (MS) is a powerful analytical approach to characterize analytes based off of the molecules' mass-to-charge ratio (m/z) and the fragmentation pattern. The functionality of mass spectrometry relies on three primary components; the ion source, mass analyzer, and detector. As the mass spectrometer measures m/z values, an ion source is required to generate charged molecules during the ionization process. Charged ions are then separated by either an electric or magnetic field, known as the mass analyzer, prior to reaching the detector. Initial characterization of an analyte is commonly based on accurate mass from the MS, but further validation can be achieved by tandem mass spectrometry (MS/MS), in which the configuration of the mass spectrometer includes an initial ion selection step followed by fragmentation and analysis of fragments, also known as daughter ions. While an in-depth description of the underlying physics that drives the instrumentation is beyond the scope of this dissertation, a brief introduction to the common techniques used in the biomedical sciences and for this dissertation is provided.

1.2.1.2 Ion Sources

The two most influential ionization techniques for the analysis of biological molecules, and thus the two ionization methods used in this dissertation, are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). Both of these ionization techniques are regarded as soft ionization techniques, meaning they induce little molecular fragmentation, unlike the ionization techniques that preceded them. The concept of electrospray ionization was

introduced by Malcolm Dole in 1968 and developed and applied for the study of large biological molecules by John Fenn in the 1980s (98,99). The basic principle of ESI is that the solvent containing analytes is infused into a charged capillary resulting in a Taylor cone formation, which emits an aerosolized spray, ultimately resulting in charged droplets that travel through an electric field towards the mass analyzer. As the solvent evaporates, the charge density of the droplets increases until the point where the charge exceeds the Rayleigh limit and the droplet undergoes Rayleigh fission to create smaller charged ions (100). The presence of gas flow assists in the process of nebulization, directs the spray towards the mass spectrometer, and aids in solvent evaporation (101). The transition of analytes to the gas phase does not require exorbitant amount energy and can proceed with minimal fragmentation. One advantage of ESI is that it ionizes molecules in the liquid phase, allowing ESI to be in-line and coupled to conventional liquid chromatography, including reverse phase chromatography (102). Additionally, the ability of the ion source to generate multiply charged ions, and thereby smaller m/z values, results in ions that are more compatible with mass spectrometers, which are often set to detect m/z values less than 2,000 Da. The presence of the same species with different charge states provides multiple opportunities to identify and fragment a molecule, which is particularly useful if the fragmentation produced by one charge state is not acceptable (103).

Franz Hillenkamp introduced MALDI as an ion source around the same time that John Fenn introduced ESI (104). In MALDI, the sample is mixed and cocrystalizes with a matrix that is present in large molar excess. Most matrices absorb UV light, which is fired in the form of a pulsed laser at the matrix crystal, causing the matrix to absorb energy from the laser, resulting in desorption of the matrix and the analyte (105). The matrix also acts as a proton donor or receptor, resulting in ionization for negative or positive ion mode, respectively (105). As the matrix absorbs the energy from the laser instead of the sample itself, MALDI is considered a soft ionization strategy and results in minimal analyte fragmentation. MALDI primarily produces

singly charged ions rendering mass spectra somewhat less complex than ESI spectra. However, multiply charged ions can be detected by altering the matrix or protocol for matrix application (106). MALDI is also more tolerant of contaminants, such as residual detergents, than ESI (107). While most studies rely on either ESI or MALDI ionization, the techniques can be quite complementary in nature. For proteomic analysis of complex samples, the use of MALDI in addition to ESI ionization resulted in 45% more protein identifications, including many that were identified based off of larger and more hydrophilic peptides, such as glycosylated peptides (108).

1.2.1.3 Mass Analyzer

Once ionized, the mass analyzer functions to separate ions based off of their m/z. Importantly, the mass analyzer is responsible for the overall resolution, mass accuracy, mass range, and fragmentation profile that can be achieved (109). The general classes of mass analyzers are Quadrupole (Q), Time of Flight (TOF), Magnetic Sector, Electrostatic Sector, Quadrupole Ion Trap (QIT), and Ion Cyclotron Resonance (ICR). Each mass analyzer relies on specific physical properties to achieve ion resolution. Quadrupole mass analyzers consist of four parallel metal rods that are used to separate ions based on their trajectory in an electric field. One pair of rods is subjected to a direct voltage, whereas the other pair has an alternating radiofrequency (RF) potential. Depending on the voltage conditions, only ions with a certain m/z will be able to reach the detector (109). The quadrupole mass analyzer has been widely adopted as it is mechanically simple and offers advantages in terms of cost, size and speed (110). The TOF mass analyzer was popular in the 1960s but was surpassed due to limitations in sensitivity, resolving power, and speed. However, instrumental improvements have overcome these limitations resulting in a renewed interest in TOF (111). The principle behind TOF is that ions with the same kinetic energy but different masses will have a velocity that is inversely proportional to the square root of their m/z value (111). Thus, the time required to navigate the flight tube will be shorter for smaller ions based off of their greater velocity. The m/z of an ion

will be calculated based on the length of time it takes the ion to navigate the flight tube and reach the detector. Similar to the quadrupole analyzer, quadrupole ion trap mass analyzers can trap and accumulate ions using an electric field for the separation of ions based off of their m/z values. Unlike the quadrupole analyzer, the quadrupole ion trap is composed of a ring electrode and two end cap electrodes. Ions are trapped in the electric field and can be ejected by increasing the RF voltage (112). The quadrupole ion trap mass analyzers provide superior sensitivity and data acquisition speeds compared to other mass analyzers but are limited by reduced resolving power and higher mass defects (107). Ion cyclotron resonance and Orbitrap mass analyzers are becoming increasingly popular, due to their higher resolving power and mass accuracy. The FT-ICR mass spectrometer uses ion optics to transfer ions into the magnetic field and the ICR cell, also known as a Penning trap. As the ions are orbiting, they are excited by an RF current to induce a current image, which can be processed by a Fourier transform (113).

Each mass analyzer offers different advantages that can be exploited depending on the nature and complexity of the experimental design. Additionally, certain analyzers seem to pair nicely with specific ion sources. The most common mass analyzers paired with a MALDI ionization source is the TOF mass analyzer (105). On the other hand, quadrupole mass analyzers have historically been interfaced with ESI sources, as quadrupoles are tolerant of high pressures, able to assess m/z values up to 3,000, and are considered low cost. More recently, quadrupole mass analyzers have been largely replaced by ion traps. FT-ICR mass analyzers can be coupled to both MALDI and ESI sources to improve resolution and mass defect errors.

As previously mentioned, tandem mass spectrometry is used to generate ion information beyond the m/z value. Collision-induced dissociation (CID) is one fragmentation technique, in which an ion is selected in the mass analyzer and fragmented through the collision of a gas in the collision cell. The triple-quadrupole mass analyzer is a common instrumental setup used to perform CID, in which a particular ion is isolated in the first quadrupole, fragmented in the

second quadrupole, and the fragments are analyzed in the third quadruple. Similarly, Q-Q-TOF instruments utilize the first quadrupole for precursor ion selection, the second quadrupole for CID, and the ToF for resolution of the product ions (107). In contrast, ion-trapping instruments, such as the QIT or FT-MS, operate in a tandem-in-time method, meaning that the fragmentation is performed in the same mass analyzer (114). In these instruments, selected ions are trapped and fragmented in the ion trap prior to the fragments being ejected onto the detector. The high MS/MS efficiencies of the ion trap instruments have made them a prominent analyzer for proteomic workflows. Other fragmentation techniques utilized in this dissertation include electron transfer dissociation (ETD) and high energy collision dissociation (HCD). For ETD, negatively charged reagent ions (fluoranthene radical anions) transfer electrons to the precursor cations causing them to fragment. For peptides, ETD fragmentation produces c and z type fragment ions of the peptide backbone (115,116). HCD fragmentation uses a beam-type CID mechanism, but operates at a higher energy than traditional CID. This generates informative low molecular weight ions that can be detected in the Orbitrap, which overcomes the one-third effect that is associated with CID in linear ion trap instruments (117).

1.2.1.4 Mass Detector

In order for a signal to be generated from the ions, mass detectors rely on the process of secondary ion amplification or the generation of a current. An electron multiplier is one of the most common detectors and relies on a series of dynodes at increasing potential. As ions strike the dynode surface, secondary electrons are emitted and attract to the next dynode, which occurs in an iterative process to allow for signal amplification (105). Eventually, the emitted electrodes are collected by a metal anode. Scintillation counters are similar to the electron multipliers in that ions initially strike a dynode, initiating a cascade of electrons, but differ in that the electrons strike a phosphorus screen that emits photons that are detected by a photomultiplier (105). The quadrupole ion trap functions to trap ions and also to detect ions. As the potentials that are

applied to the electrodes of the ion trap are altered, ions reach an unstable trajectory and are scalned out of the ion trap and are monitored by an external detector (112). Charge detectors, which detect a current induced by a passing ion, are common in FT-MS instruments. In the FT-ICR-MS, the m/z of an ion is determined by its ion cyclotron resonance frequency in the magnetic field. The m/z in the Orbitrap mass spectrometer is determined by its oscillation frequency in an electric field (118). This process generates an image current of an ion, which can be Fourier transformed to ultimately yield the m/z.

1.2.2 Mass Spectrometry Applications in the Biomedical Sciences

Early applications for mass spectrometry were predominantly related to physics based research, but the introduction of MALDI and ESI ionization sources in the 1980s ushered in a new wave of biomedical research using mass spectrometry (119). These soft ionization sources enabled the analysis of large molecules instead of focusing on molecules less than 1kDa. The era of genomics enhanced the desire to assess broad “omic” studies of specimens in general, or for the study of diseases. As the mass spectrometer is able to simultaneously detect and characterize analytes in a single experiment as a function of their m/z and downstream fragmentation, the technique is well suited for "omics" research, in which thousands of analytes may be present. To this end, mass spectrometry has been instrumental in lipidomics, metabolomics, proteomics, and the global analysis of post-translational modifications (PTMs) (120–124).

Applications relating to general proteomic experiments will be outlined in this section, with a more in-depth review of glycoproteomics presented later in the dissertation. Proteomics is the global study of proteins and their modifications in a sample. A distinction can be drawn between classical proteomics, which aims to identify the entire proteome; and targeted proteomics, which assesses a defined subset of the proteome (i.e. glycoproteins) (103). In addition to protein identification, proteomic analysis can also be used to interrogate protein abundance in a quantitative fashion. The standard workflow typically involves sample

fractionation prior to analysis by mass spectrometry, which functions to allow the detection of lower abundance ions. This is required, especially for complex biological fluids, due to the high dynamic range of proteins in the sample. Proteomics is typically done at the peptide level (bottom-up proteomics), further complicating the sample complexity by producing multiple peptides for a single protein. The detection of peptides by the mass spectrometer is typically accomplished in positive ion mode, where the peptides are analyzed at a low pH where arginine, lysine and histidine amino acids are protonated (103). Fractionation and chromatography can be done at the protein level, through techniques such as SDS-PAGE, or at the peptide level using SPE cartridges. More commonly, chromatography can be applied in an on-line fashion by connecting chromatography instruments (i.e. HPLC) directly to mass spectrometers.

The detection of product ions only provides a list of m/z values, so fragmentation of each ion is required to determine the peptide sequence. Chromatographic separation prior to MS analysis allows concentration and resolution of peptides prior to mass spectrometery, resulting in more protein identifications. The selection of which ions to fragment can proceed through a variety of different mechanisms which ultimately play a role in the outcome of the experiment. The most common method is data-dependent acquisition (DDA) workflows, in which the most abundant ions throughout the chromatographic separation are selected for fragmentation (125). While commonly employed, DDA suffers in its ability to characterize low-abundance peptides in complex biological mixtures (125). To compensate for these limitations, several alterations to conventional DDA workflows can be applied (126). These include targeted LC-MS strategies, such as selected reaction monitoring (SRM) and parallel reaction monitoring (PRM). These targeted approaches function by acquiring MS/MS spectra for MS ions of interest, even when they are not the most abundant ions present (125). Alternatively, data-independent acquisition (DIA) strategies, fragment all m/z values, regardless of ion intensity, by fragmenting across m/z ranges (127).

While the ultimate goal of proteomics is to characterize the entire proteome of a sample, this remains a significant challenge. Recent successes rely upon advances in instrumentation and data processing strategies. During the last decade, mass spectrometers have made significant improvements in resolution, mass accuracy, sensitivity and scan speed (128). For shotgun proteomics, in which a peptide digest is processed by LC-MS/MS, peptide identification relies on database searching algorithms for protein identification. The software works to compare tandem mass spectra from the experimental data to that generated in an *in silico* protein digest, to determine what peptides are present in the sample, to make inferences about the proteins present. Many software packages are available to process complex shotgun proteomics data including ProteoIQ, Proteome Discover, Scaffold, and MaxQuant. Improvements in instrumentation and data processing software have resulted in a rapid growth of proteomics as a field, but improvements can still be made. The human proteome (~100,000 proteins) vastly exceeds the genome (~30,000 genes), even without considering the additional complexity provided by PTMs, mutations and truncations. Considering the added diversity, it is not unreasonable to expect the proteome to exceed 2,000,000 proteins (103,129).

1.2.3 Mass Spectrometry Imaging

Imaging Mass Spectrometry (IMS) is recent addition to the field of mass spectrometry. This technique enables the simultaneous detection of analytes while retaining their spatial localization within tissue sections. General "omics" techniques reflect trends in analyte expression across the entire tissue, but may be biased by the presence of differing histology. To compensate for tissue heterogeneity, laser capture microdissection (LCM) is one approach used to selectively enrich specific cells of interest (130,131). However, the utility of LCM is limited in that the approach is time consuming and requires a high number of cells. Another approach to assess analyte distribution across a tissue section relies on voxelation followed by LC-MS, in which an individual analysis is done for each defined region, or voxel, of the tissue (132). While a

large number of analytes are detected, the spatial resolution of this approach is limited by the size of the voxels and length of time required, as sample preparation and data analysis must be performed on each voxel individually. To compensate for these limitations, IMS has been adapted for use with secondary ion mass spectrometry (SIMS), MALDI, and desorption electrospray ionization (DESI). Although SIMS-IMS is capable of achieving a higher overall spatial resolution (as small as 50 nm), the approach is not commonly used because of undesired fragmentation attributed to the high-energy ions used to promote ionization (133). Recent sample preparation strategies have improved the efficiency of SIMS ionization, which has a niche utility in imaging at the single cell level (134). DESI-IMS, in which ions are desorbed by targeting charged droplets and ions at the surface of the tissue, suffers from low spatial resolution and low sensitivity (133). MALDI-IMS is capable of achieving a relatively high spatial resolution, while inducing less fragmentation compared to SIMS, and thus represents the most popular ionization source for IMS analysis. Efforts to optimize the sample preparation are essential to MALDI-IMS experiments, as uneven distribution of matrix or the formation of large matrix crystals will negatively impact the experiment. Importantly, all IMS platforms generate parent ions that can be fragmented for further structural confirmation. As a platform, MALDI-IMS has been instrumental in the analysis of lipids, proteins, peptides, therapeutics, and metabolites (135–143).

All MALDI-IMS experiments require the same basic sample preparation steps prior to imaging, although variations to the general workflow can significantly impact the overall results or class of biological molecule that is detected. In the workflow schematic provided (Figure 2), solid connecting arrows represent conserved steps across all imaging experiments, while steps connected by dashed arrows can be utilized for selected experiments. Initially, tissues are sectioned and placed on a platform compatible with a MALDI adaptor, most commonly glass slides. A homogenous layer of matrix is then applied to the slide containing the sectioned tissue and allowed to crystallize. During the imaging process, a MALDI spectra is generated at specified

increments across the tissue section (termed spatial resolution). Software for data processing merges all of the individual spectra while maintaining their original location in the tissue section. Individual m/z values can be selected and visualized by the relative intensity for the ion across the entire tissue section. This approach is specifically useful in discovery phase research, where undefined ions are only observed in a specific compartment of a tissue, but are not commonly observed in global omics research due to ion suppression from more abundant species. While the general approach applies to all MALDI-IMS experiments, efforts to optimize sample preparation have been performed to enhance sensitivity performance for certain molecular targets. Initial MALDI-IMS workflows utilized fresh/frozen tissue sections, but the adaptation of immunohistochemical workflows has enabled MALDI-IMS analysis of formalin-fixed paraffin-embedded (FFPE) tissue blocks and tissue microarrays (144). Depending on the target of the imaging experiments, different washing procedures can be employed. While MALDI-IMS of lipids does not require washing of the slide, washing with ammonium formate or ammonium acetate can improve the overall ionization efficiency and number of ions observed (145). To analyze proteins or peptides by MALDI-IMS, lipids should be removed by washing the slide in organic solvents to mitigate ion suppression (144,146). Following the wash, *in situ* enzymatic digestion can be performed by the application of an enzyme (i.e. trypsin) followed by MALDI-IMS. Both the method of matrix application and type of matrix used can significantly affect the type of molecules detected and the overall quality of the experiment. Automated spray stations and microdroplet inkjet printers are among the most common instruments used for matrix application, but suffer from solvent-induced analyte delocalization and the formation of large matrix crystals (138,147). On the other hand, matrix application by sublimation produces a homogeneous layer matrix with reduced sizes of matrix crystals (136). Different matrix compositions can improve analysis of specific classes of molecules in MALDI-IMS, largely reflective of traditional MALDI-MS. Both 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-

hydroxycinnamic acid (CHCA) are commonly applied matrices for MALDI-IMS of lipids, but are primarily used for the detection of PCs and SMs in positive ion mode. 1,5-Diaminonaphthalene (DAN) matrix was demonstrated to be efficient in positive and negative ion mode, enabling the detection of lipids that are primarily detected in negative ion mode, such as sulfatides, ceramides and gangliosides (136).

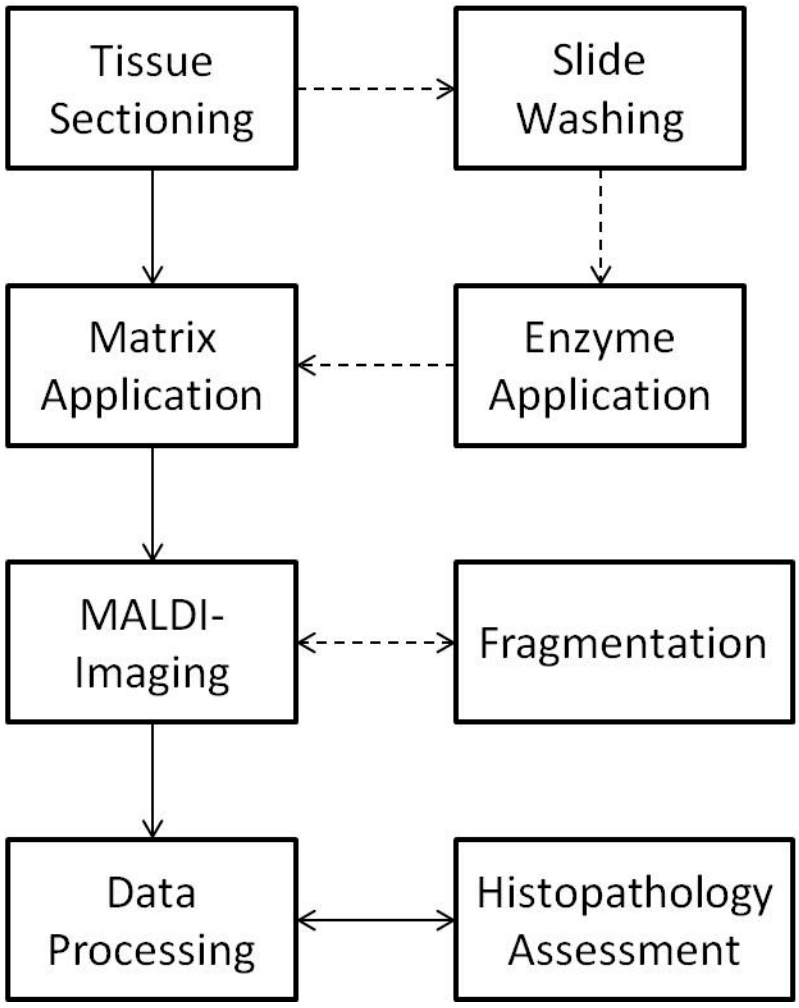


Figure 2. MALDI-IMS Workflow Schematic. Universal (Solid Arrows) or optional (dashed arrows) steps can be optimized to achive better MALDI-IMS results. Optimal conditions are often dependent on the tissue and experiment performed.

1.2.4 Mass Spectrometers in the Dissertations

1.2.4.1 Bruker SolariX FTICR

A majority of the MALDI MS and MALDI-IMS experiments in this dissertation were carried out on Bruker SolariX 7 Tesla FTICR mass spectrometer. The instrument possesses both an ESI and MALDI ion source. After ionization in the MALDI source by a smartbeam-II laser, ions are focused into an ion beam using DC and RF voltages. At this stage, ions pass through a vacuum stage and are decelerated. Ions then traverse to the quadrupole, in which a specific m/z window can be defined to only allow ions within that range to pass through to the hexapole. If desired, specific ions can be fragmented in the collision cell. Ions are then focused by optical lenses prior to entering the magnetic field of the ICR cell, which is under ultrahigh vacuum conditions. The infinity cell, which exists at the center of the 7T magnet, consists of three sets of electrodes that are used for ion trapping, excitation and detection. The magnetic field traps ions and induces cyclotron motion, but a small electrical field is required to prevent ions from escaping the ICR cell. The trapped ions are then excited by a RF frequency sweep, which increases the radius of selected ions and enables detection by the detection plate (148).

1.2.4.2 Thermo Orbitrap Elite

The Thermo Scientific Orbitrap Elite mass spectrometer is considered a hybrid mass spectrometer that couples a linear ion trap with an Orbitrap mass analyzer. The linear ion trap on the front-end is the Thermo Scientific Velos Pro mass spectrometer. This functionality provides excellent sensitivity and data acquisition speed associated with ion trap mass analyzers, and is capable of trapping, isolating and fragmenting ions (149). The linear ion trap possesses an independent MS detector, so that the ions can be detected in the ion trap or proceed to the Orbitrap for further analysis. The transition to the Orbitrap requires ions to move from the RF-

only octopole into the curved linear trap (C-trap). As the ions go from a gas-free environment to the gas-filled C-trap, ions lose their kinetic energy by colliding with the nitrogen collision gas. Ions are transferred to the Orbitrap by ramping off RF voltages applied to the C-trap and applying extracting voltages onto the Orbitrap electrodes. As described in Thermo literature, the Orbitrap analyzer consists of a central spindle-shaped electrode surrounded by bell-shaped outer electrodes. Injected ions begin axial oscillations around the central electrode of the Orbitrap, and applied voltages prevent collision with the electrode. An image current is detected by two split halves of the outer electrode of the Orbitrap, which can ultimately determine the frequency of axial oscillation and m/z ratio of ions. The Orbitrap mass analyzer enables the detection of analytes with high mass accuracy, resolution and expands the dynamic range that can be detected (149). The combination of two divergent mass analyzers enables peptides to be measured at sub-ppm mass accuracy which significantly improves bottom-up protein identification. The Orbitrap Elite is setup to allow CID, HCD, and ETD fragmentation of selected precursor ions (116). For HCD fragmentation, ions enter the HCD cell from the C-trap and the fragment spectra are detected in the Orbitrap analyzer. The ETD source is located on the back of the Orbitrap Elite mass spectrometer and can fragment peptide ions in the linear ion trap by the transfer of electrons from negatively charged reagent ions.

1.3 Glycosylation

1.3.1 Broad Overview of Glycosylation

1.3.1.1 The Complexity of Protein Glycosylation

Glycosylation refers to the addition of monosaccharide or oligosaccharide chains, commonly referred to as glycans, onto biomolecules such as proteins, lipids, proteoglycans and glycophosphatidylinositol (GPI) anchored proteins. Recently, free, unbound glycans have been discovered and implicated as important signaling molecules (150). It is estimated that over 50%

of human proteins are glycosylated, making it the most common and complex post-translational modifications. There are nine monosaccharide units from which glycans are constructed, but additional diversity can be achieved by further modification of those monosaccharides (150). A general classification system groups monosaccharides based on their molecular weight. Hexose (Hex) monosaccharides consist of glucose (Glu), galactose (Gal) and mannose (Man) residues. N-acetylhexosamine (HexNAc) monosaccharides, which differ from Hex monosaccharides by the addition of an acetyl functional group, consist of both N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc). The remaining monosaccharides include fucose (dHex, fuc), N-acetylneuraminic acid (NeuAc), glucuronic acid (GlcA) and xylose (Xyl). Figure 3 provides a schematic of monosaccharides compositions as they will appear throughout the rest of this dissertation. While most of these monosaccharides are present throughout all species and across all glycoconjugates, some monosaccharides display a specificity to certain biological species or types of glycoconjugates. In contrast to plants, in which complex N-glycans contain xylose residues, these residues are only observed in humans on proteoglycans (150,151). In addition to variation in the type of monosaccharides found, species wide differences in glycosylation also appear as differences in glycosidic bond linkages. For example, plant N-glycans express non-human glycan linkages, such as α1,3-linked core fucose residues (151). While limited in the number of monosaccharides, glycosylation is quite complex. Unlike transcription and translation, the process of glycosylation is not template encoded but depends on a variety of cellular factors. Substrate availability, developmental stage, nutrient availability, pH and enzyme localization all alter the observed cellular glycoforms, and allow for the glycome to respond to subtle biological changes (152). Glycosyltransferases and glycosidases, which add and remove monosaccharides from glycans, are responsible for the structure of the glycoform. It is estimated that around 5% of the genome encodes enzymes that are involved in cellular glycosylation in some fashion (150). Multiple enzymes can reside in the same region of the cell and compete over the same substrate,

resulting in a complex array of glycoforms (152). However, glycan synthesis is not completely random, as glycosyltransferases exhibit some degree of specificity for the protein, glycopeptide and growing glycan structure (152). Glycosidases promote trimming of the glycan, both during or after synthesis, but are equally responsible in yielding glycan structures. As these enzymes function sequentially, the absence of a single glycosyltransferase that acts early in glycan synthesis can result in pronounced differences in all observed glycoforms (153,154).

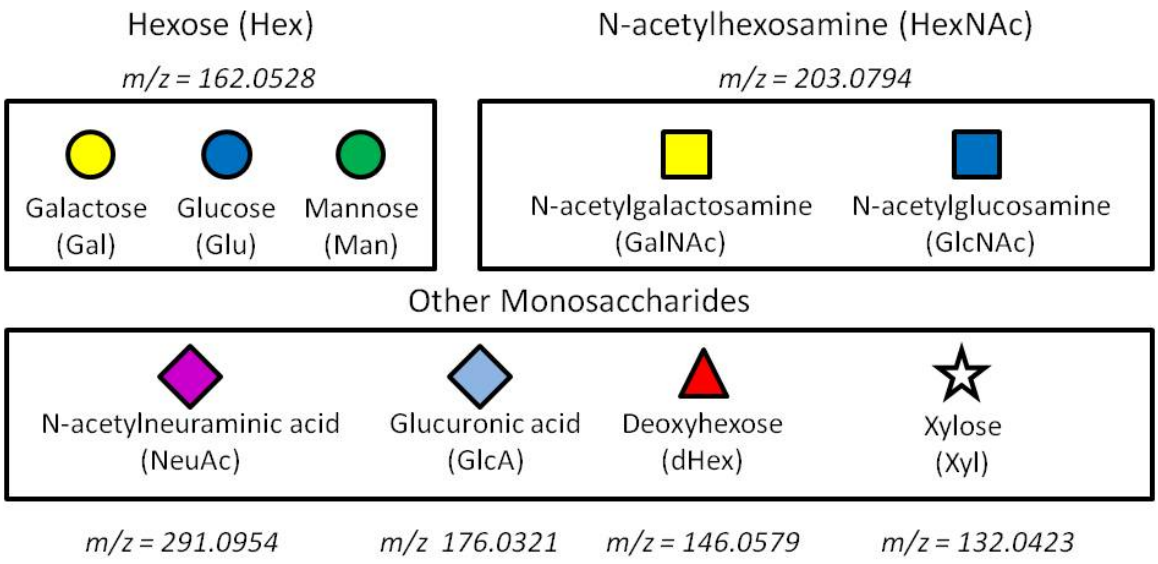


Figure 3. Molecular Weight of Monosaccharides.

Glycosylation, like all post-translational modifications, renders the proteome more complex than the transcriptome would suggest. Given the seemingly stochastic nature of glycan synthesis, it is not surprising that glycosylation is associated with an enormous degree of heterogeneity. Heterogeneity in protein glycosylation can be further classified as microheterogeneity and macroheterogeneity. In the field of glycobiology, microheterogeneity refers to the observation that multiple glycoform compositions are present at an individual glycosylation sites. Furthermore, glycan branching results in divergent glycans with the same monosaccharide composition (155). It is estimated that the human serum N-glycome consists of

around 300 to 500 glycan compositions, representing close to 5000 structures (156). Macroheterogeneity of glycosylation refers to the rate of glycan occupancy on an individual site of glycosylation. Even with the complex nature of cellular glycan processing, many proteins display consistent glycosylation patterns across independent samples, suggesting an important role of glycosylation in cellular function.

1.3.1.2 Physiological and Pathological Roles of Glycosylation

Glycosylation is integral to cell survival, demonstrated by role of glycoconjugates in cell viability and embryonic development. While specific changes in glycosylation are well documented, the global elimination of a single class of glycan results in cell death (157). More subtle modifications of the glycan biosynthetic pathway have variable impacts on cell viability, depending on what stage of glycosylation is impacted. For N-glycan synthesis specifically, termination of N-glycan synthesis prior to the formation of hybrid and complex glycans is embryonic lethal, while modification of terminal glycan structures elicits more limited abnormalities (153). The role glycosylation plays in embryonic development is emphasized by the detection of glycans which are expressed embryonically but are absent in adult tissues (152). Even the process of fertilization is mediated by glycan epitopes on the sperm interacting with selectins on the ECM of the oocyte (158).

The biological role of glycosylation is not limited to the developmental stages of cells. Glycosylation regulates normal biological activity through direct and indirect effects. Direct effects refer to the interaction of the glycan with various glycan binding proteins (GBPs), including lectins, GAG-binding proteins and antibodies (152). Biologically, direct interactions are important in the cells interaction with cells, the ECM and other molecules. While all GBPs use direct effects to modulate biological function, the mechanisms for binding vary between the types of GBPs. For instance, lectins recognize terminal glycan epitopes and fit them into shallow binding pockets, while GAG-binding proteins use electrostatic interactions between the anionic

GAG chains and cationic amino acids (150). While direct effects typically lack specificity between similar glycoforms, the interaction can sometimes be highly specific. For example, P-selectin exclusively recognizes and binds to P-selectin glycoprotein ligand-1 (PSGL1) when SLeX is expressed on a core 2 O-glycan, but not on a core 1 O-glycan (159). While direct effects depend on the glycan itself, indirect effects of protein glycosylation refer to effects that are caused by the physiochemical properties of the glycan or alterations in the properties of the protein resulting from the addition of bulky hydrophilic sugar group. These indirect effects impart broad regulatory functions on both secreted and membrane proteins. Glycosylation of secreted proteins reduces nonspecific interactions, protects the protein from proteases and antibodies, and improves the solubility of the secreted proteins (160). These indirect effects may or may not depend on the composition of the glycan. Terminal sialic acid residues are important for reducing nonspecific interactions due to their negative charge, while the general presence of bulky, hydrophilic groups can shield sites from protease digestion (160). In addition to these effects, indirect effects of membrane protein glycosylation ensure proper folding of the protein, enable ECM localization and maintain the tissue structure and integrity (150). The fact that all cells are coated with an external layer of glycoconjugates, highlights the importance of glycosylation in molecular recognition and signaling. In fact, glycosylation can mediate downstream signaling by directly modulating functionality or cell surface receptor expression. For example, both glycosylated and deglycosylated beta-human chorionic gonadotropin exhibit similar receptor binding affinity, but only the glycosylated hormone can stimulate adenylate cyclase (150). Glycosylation of epidermal growth factor receptors (EGRFs) has been directly linked to elevated signaling due to reduced receptor turnover and prolonged exposure on the membrane (161).

Glycosylation plays important roles in regulating the immune system and inflammatory responses. In the immune system, the presence of protein glycoforms can be used to define self glycans but also impacts the efficacy of immunoglobulins. Human ABO blood groups are

determined by the status of specific glycosyltransferases possessed by an individual, resulting in specific oligosaccharide moieties on proteins and lipids (162). For this reason, when blood is not appropriately matched during transfusions, antibodies recognize foreign glycan antigens and elicit an immune response. Autoimmune diseases, such as lupus or rheumatoid arthritis, arise due to the aberrant immune response triggered by self-glycans (157). In mice, the absence alpha-mannosidase-II (αM-II) mimics the systemic lupus erythematosus phenotype. Functionally, the absence of αM-II results in incomplete N-glycan branching and an increase in hybrid-type N-glycans with a terminal mannose, which results in an immune response and inflammation (163). The glycosylation patterns of immunoglobulins also plays a role in regulating the efficacy of the mounted immune response. The absence of fucose or sialic acid residues on IgG enhances the antibody dependent cell mediated cytotoxicity (ADCC) as a function of higher-affinity binding to FCγRs (164). IgG glycosylation can elicit pro- or anti-inflammatory effects. Patients with rheumatoid arthritis frequently present with reduced galactosylation levels of serum IgG, even prior to the onset of RA, suggesting that the pro-inflammatory effect of aberrant glycosylation ultimately drives rheumatoid arthritis (164). Given the role glycosylation plays in regulating basic immune response functions, it is unsurprising that defective immune cell function is the most common phenotype in mice that survive deletion of specific glycosyltransferases (157). The regulator role of immunoglobulin glycosylation on inflammation is not the only mechanism by which inflammation is linked to glycosylation. Cytokines generated in response to pathogen-associated molecular patterns (PAMPs) or damage-associated molecular pattern molecules (DAMPs), interact with glycoprotein receptors and lectins during leukocyte homing to inflamed tissues. Glycan fragments released during the process of leukocyte degradation of ECM glycoconjugates signal for more leukocyte chemotaxis and activation, resulting in further inflammation (157).

In addition to normal biological functions, abnormal glycosylation has been observed in nearly every disease. Diseases may be caused by the presence of unique glycans, elevated expression of glycans, or the absence of glycans in a biological system. In cardiovascular disease, evidence suggests that glycosaminoglycans (GAGs) on proteoglycans are responsible for initiating early stages of atherogenesis (165). After initiation, monocyte entry into the subendothelial regions of blood vessels relies upon the interaction of SLeX on PSGL-1 with endothelium selectins. This observation was verified in vivo by the reduction of fatty streaks in mouse models that are deficient in P- or E-selectins compared to control mice (166,167). In contrast, the absence of specific glycans can result in disease. One such example is inclusion cell disease, a lysosomal storage disease that is characterized by the buildup of substances in the lysosome. This buildup is caused by the absence of enzymes in the lysosome due to the absence of mannose-6-phosphate, which targets enzymes to the lysosome (153).

Host-pathogen interactions are often mediated by the interaction of pathogens with specific glycan patterns on the host cell. The functionality of the influenza virus depends on two glycoproteins, hemagglutinin and neuraminidase. Hemagglutinin binds to sialic acids on the host-cell glycoproteins and glycolipids to initiate viral infection. Following viral replication, neuraminidase cleaves the sialic acid to allow the virus to release from the host-cell and continue to infect more cells (168). Similarly, cholera toxin binds sialylated glycolipids, primarily GM1, present on the intestinal epithelia (169). Occasionally, the absence of specific glycoforms can result in pathogen invasion. The absence of terminal α-GlcNAc residues on mucin core 2 O-glycans results in Helicobactor pylori colonization (170). Apart from the direct host-pathogen interaction, glycans on the viral glycoprotein coats can aid in pathogenesis. In human immunodeficiency virus (HIV), N-glycosylation on the envelope glycoprotein (Env) provides resistance to antibody neutralization and enhances infectivity (171). Alternatively, pathogens can adapt to mimic host glycan structures in an effort to evade the hosts immune system (172).

Given the broad nature of glycosylation in diseases, it is understandable that glycosylation is frequently exploited for therapeutic purposes. Rational vaccine design has benefited from an increased knowledge of aberrant glycosylation in diseases and improved synthetic approaches for glycopeptide synthesis. Traditionally, carbohydrate antigens initiate low-affinity, short-lived IgM responses, but fail to induce a T-cell response (173). However, by combining carbohydrate antigens to a specific carrier protein, T-cell responses can be initiated. This approach has been applied in the generation of anticancer vaccines, which aim to eliminate preexisting cancer cells (172). Using this rationale, Muc1 glycopeptides were determined to initiate immune responses and antibody-dependent cellular cytotoxicity (ADCC) more efficiently than non-glycosylated Muc1 peptides (174,175). Neu5Gc-bearing gangliosides have been successful targets in immunotherapy based on the antigenicity of the non-self Neu5Gc monosaccharide (172). Host-pathogen interactions are another area where knowledge of glycosylation can be used to improve therapeutic outcomes. As the initial stage of pathogen infection often involves the pathogen interaction with specific glycan sequences on the target cells, the disruption of this interaction is associated with improved therapeutic effects. The use of free glycans that mimic the pathogen interaction sequence has been reported to inhibit infection by various pathogens (172). Free glycans, along with the presence of immunoglobulins, are the suggested reasons that breastfed are on average healthier than formula fed infants (176,177). In addition to disrupting the initial interaction between the host and pathogen, the glycosylation related mechanisms can also be exploited to inhibit the release of the pathogen from cells. An example of this is the influenza virus, which requires neuraminidase activity to release from the cell surface of infected cells. Two current treatments for the influenza virus, zanamivir and oseltamivir, function by inhibiting neuraminidase activity and thereby preventing the release and further propagation of the virus (178,179). Unique glycosylation patterns on viral coats are also being exploited for rational vaccine design. The HIV coat protein is heavily glycosylated, but

with oligomannose-type glycans that are rare in cell surface and secreted glycoproteins (180). Antibodies that recognize various glycan-peptide epitopes on the HIV coat protein are projected to be useful in developing HIV vaccinations. A therapeutic approach for treating patients with glycosylation-based genetic disorders is to provide exogenous monosaccharides. Patients diagnosed with leukocyte adhesion deficiency II (LAD II), a disease associated with a loss of SLeX-mediated leukocyte rolling, display improved immune function following treatment with exogenous L-fucose (181). Conversely, lysosomal storage diseases, characterized by defective enzyme activity of lysosomal enzymes, can be treated by enzyme replacement therapy. For instance, both Fabry disease and Mucopolysacchridosis I, can be treated by the administration of recombinant enzymes α-galactosidase A and α-L-iduronidase, respectively (182,183).

Biotherapeutic glycoproteins represent a robust market for pharmaceutical companies. The presence of specific glycoforms on glycoprotein therapeutics can modulate immune response, impact therapeutic pharmacokinetics, and therapeutic functionality (184). Therapeutic glycoproteins are often expressed in nonhuman cell lines and are therefore subject to non-human epitopes and linkages. Efforts to humanize glycans are being investigated through the use of cell glycosylation mutants and transgenically altered cell lines. By using cell lines with specific glycosylation machinery, scientists can exert some control over the ultimate glycoforms on therapeutic glycoproteins, which play important roles in therapeutic efficacy. Erythropoietin is a prime example of this, where deglycosylated erythropoietin exhibits similar in vitro activity compared to glycosylated erythropoietin, but the in vivo activity is reduced by about 90%. The reduced in vivo activity was found to be a function of rapid renal clearance of the deglycosylated erythropoietin (150). Similarly, proper glycosylation of antibodies also increases the circulatory lifetime (185).The ability to regulate the sites and structures of glycoforms is a focal point in glycoprotein therapeutic research.

1.3.1.3 Mechanisms Involved in Protein Glycosylation

The addition of N-glycans to proteins occurs co-translationally in the endoplasmic reticulum and is followed by further glycan processing in the Golgi. During protein synthesis, an amino acid sequence targets proteins to the secretory pathway resulting in their co-translational transportation to the ER, where a signal-recognition particle (SRP) binds the receptor in the ER membrane, promoting the protein complex to enter the ER through the translocon (186). The en bloc addition of N-glycans is essential for secretory and membrane protein folding. Two lectin-like chaperones in the ER, calnexin (CNX) and calreticulin (CRT), bind glucose residues on the nascent glycan to ensure proper folding and prevent aggregation of the protein (187). Ultimately, glycosidase trimming releases properly folded proteins from the CNX/CRT complex. Misfolded proteins are reglucosylated by UDP-Glc:glycoprotein glucosyltransferase (UGGT), which can then interact with CNX/CRT or be transported to the cytosol and degraded by ER-associated degradation (150). Properly folded proteins are packaged in COPII-coated vesicles and then transported to the Golgi where further glycan processing occurs.

N-glycan synthesis begins with the transfer GlcNAc-P from UDP-GlcNAc to the lipid-like precursor, dolichol phosphate, to form dolichol pyrophosphate N-acetylglucosamine (Dol-P-P-GlcNAc). Dol-P-P-GlcNAc is further extended on the cytoplasmic face of the ER membrane by the sequential addition of one more GlcNAc and five mannose residues. A "flippase" is then used to flip the N-glycan precursor to the lumen of the ER, although the mechanism is not completely understood. In the ER lumen, the glycan is further extended by the addition of four mannose residues and three glucose residues, completing the synthesis of the mature N-glycan precursor, Glc3Man9GlcNAc2-P-P-Dol. The transfer of activated sugars to the growing N-glycan is mediated by the *ALG* genes, which encode unique glycosyltransferases that carry out specific transfers of sugar residues at defined stages of glycan synthesis (188,189). The completed precursor (Glc3Man9GlcNAc2-P-P-Dol) is then transferred to an asparagine residue in the N-

glycan consensus sequence by an oligosaccharyltransferase (OST) complex. The N-glycan consensus sequence is N-X-S/T, where X is anything other than proline. After protein glycosylation, the glycan is further trimmed in the ER by α-glucosidases I and II and α-mannosidase I (Figure 4a). Trimming resumes in the *cis*-Golgi, where glycans can be processed to Man5GlcNAc2, which serves as the precursor to complex and hybrid glycan structures. Glycans that are not processed or incompletely processed by α1-2 mannosidases in the *cis*-Golgi cannot be further modified and are detected as high-mannose glycans (Man5-9GlcNAc2) (Figure 4b). In the *medial*-Golgi, Man5GlcNAc2 is acted upon by N-acetylglucosaminyltransferase I (GlcNAcT-I) and α-mannosidase II to form GlcNAcMan3GlcNAc2. Decoration of the GlcNAc residue at this stage results in the formation of hybrid N-glycans (Figure 4c). Alternatively, the transfer of an additional GlcNAc by GlcNAcT-II generates the base of complex N-linked glycans. Triantennary and tetraantennary complex glycans utilize GlcNAcT-IV and GlcNAcT-V to add additional branches, while GlcNAcT-III transfers a bisecting GlcNAc residue onto the complex glycan (Figure 4d). Finally, in the *trans*-Golgi, further maturation may take place to provide additional glycan diversity. Among the most common additions are the transfer of β-linked galactose residues to the non-reducing end, the addition of α1,6-linked fucose residues to the GlcNAc directly bound to the asparagine (termed core fucosylation), and the addition of sialic acids and fucose residues on branched LacNAc chains.

Secreted and membrane proteins are frequently O-glycosylated with mucin-type O-glycans, which can appear as linear or branched arrangements of glycans with an O-GalNAc at the reducing end. While this type of glycosylation is termed mucin-type, it is also present on proteins that are not part of the mucin family. In fact, it is estimated that over 80% of proteins that are processed through the secretory pathway are O-glycosylated (190). Unlike with N-glycosylation, O-glycosylation does not appear to occur on conserved consensus sequences, but does target serine or threonine amino acids. The synthesis of mucin-type O-glycans is initiated in

the golgi by the transfer of GalNAc from UDP-GalNAc to a Ser/Thr in an α-linked manner by the polypeptide GalNAc-transferase (ppGalNAcTs) family of enzymes, forming the Tn antigen (152). The Tn antigen is further modified to form core 1-4 structures, and less frequently core 5-8 structures. The core 1 structure (T antigen) is synthesized by the transfer of Gal from UDP-Gal to Tn by the enzyme Core 1 β3-galactosyltransferase (C1GalT1, T-synthase) in the *cis* and *medial* Golgi. Interestingly, this enzyme is associated with a molecular chaperone Cosmc, the only known glycosyltransferase chaperone. Cosmc functions to insure proper folding of T-synthase in the rough ER (191). Alternatively, Core 3 N-acetylglucosaminylfransferase (C3GnT) transfers a GlcNAc from UDP-GlcNAc in a β1,3 fashion to form Core 3 structures. The family of Core 2 GlcNAc Transferases (C2GnT) or a subset of the family, the C2/4GnT enzyme, add a branched β1,6-linked GlcNAc to Core 1 and Core 3 glycans respectively, generating Core 2 and Core 4 O-glycans. The core can be further extended to generate biologically relevant glycan moieties, such as blood group antigens (ABO), Lewis antigens, and large extended structures.

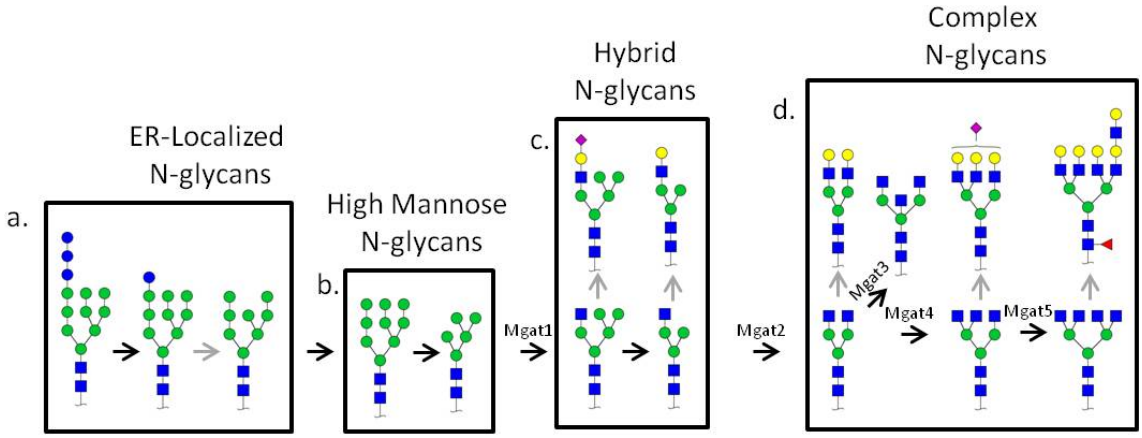


Figure 4. N-Glycan Processing. The glycan precursor is transferred onto proteins in the ER and upon trimming of glucose residues, ensures proper folding (a). The glycoprotein is then transferred to the Golgi, where high mannose glycans are formed in the *cis*-Golgi (b), hybrid and complex glycans are formed in the *medial*-Golgi (c,d), and terminal modifications are added in the *trans*-Golgi.

Nuclear and cytoplasmic proteins may also be glycosylated to include O-GlcNAc, O-Glc or O-Fuc modifications. O-GlcNAc is an especially dynamic modification that can crosstalk with phosphorylation to regulate signal transduction (152). O-GlcNAc cycling by O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA) can rapidly add or remove the modification from protein substrates to respond quickly to biological signals (192). Ultimately, the process of O-GlcNAc glycosylation allows for the timely regulation of cell cycle progression, response to stressors, and cell survival (192).

1.3.1.4 Glycosylation in Cancer

The first example of aberrant glycosylation in cancer was discovered in the 1950s by the observation that peptides were larger in transformed cells compared to untransformed cells. It is now known that this increase in peptide molecular weight is linked to elevated levels of β1,6 branched N-glycans. Around the same time, plant lectins demonstrated enhanced binding to tumor cells compared to non-tumor cells (193,194). In the 1970s and 1980s, following the discovery of monoclonal antibodies, researchers began to immunize mice against tumor cell extracts to generate antibodies. A majority of the antibodies generated act directly against a glycan antigen or glycopeptide epitopes, including B72.3, CA19-9, CA15-3, and CA-125 (195–199). Since these original experiments, alterations in glycosylation are now recognized to be a hallmark of cancer.

Cell surface changes in glycosylation can be categorized as oncofetal antigens, neo-antigens, and normal antigens. Oncofetal antigens are glycan epitopes that are expressed during embryonic development, are absent in normal adult tissue, but reappear in the presence of tumor tissue. In contrast, neo-antigens are novel glycan structures that are not present in embryonic or adult tissues. Oncofetal antigens and neo-antigens primarily exist in the form of truncated O-glycans (152). Both structures make ideal disease markers as they are tumor specific and rarely detected in patients with benign disease. Tn, sialyl-Tn and T antigens have been classified as pan-

carcinoma antigens based off of their presence in most cancers. In fact, T and Tn antigens have been observed in around 90% of carcinomas, while the two antigens are rarely co-expressed in benign diseases (200). As these truncated antigens are often co-expressed in tumors, comparing the ratio of Tn to T antigen can provide prognostic information regarding cancer progression (201). Normal antigens are glycans that are present in normal physiological situations but have altered levels of expression that correlate with a disease. Normal antigens most often vary in their terminal and extended structures. Unlike oncofetal and neo-antigens, alterations in normal antigens exist on N-linked glycans in addition to O-linked glycans. While the form of aberrant glycosylation varies across different forms of cancer, the general trends of elevated β1,6 GlcNAc branching, decreased β1,4 bisecting GlcNAc, increased cell surface sialylation, elevated core fucosylation, elevated expression of Lewis antigens, and alterations in ABO structures are commonly detected (158,198,202–206). An example of system dependent variation in glycan expression is the deregulation of ABO structures in cancer. ABO structures are not observed in healthy colon tissues but are present in tumor colon tissue. In contrast, lung cancer tissues exhibit mismatched ABO structures compared to the normal tissue (206–208). Another example of the differential effect of glycosylation is the role of bisecting GlcNAc in reducing tumor volume and metastasis in most cancers, but promoting liver metastasis in melanoma and leukemia cell lines (209). The identification of aberrantly expressed glycans can often be linked to the glycosyltransferases that are responsible for their synthesis (210,211).

The mechanisms employed by glycosylation to regulate tumor initiation to metastatic progression are still under investigation. Truncated O-glycans are detected early during malignant progression and are thus hypothesized to be important to the early stages of tumor development (198,212–214). Other modifications, such N-glycan branching, present later during the tumorigenic process and are mainly linked to the later stages of tumor growth and metastasis. While it has long been suggested that interactions of branched glycans with GBPs promotes

metastasis, a recent manuscript has hypothesized N-glycan branching may also alter the half-life and conformation of cell surface receptors (215). Changes in terminal monosaccharide residues can impact the functional role of the glycan. Many tumors, including those of the pancreas, express elevated levels of sialyl Lewis antigens, which have terminal sialic acid residues. These antigens play important roles in metastasis through their interaction with selectins. E-selectins enable tumor cells in the vasculature system to adhere and infiltrate distant sites, while P-selectins interact with the cancer cell and protect it from the immune system (152). It is through this process that cimetidine and heparin, which have been reported to block the interaction of sialyl Lewis antigens with selectins, are thought to improve cancer-related disease survival (152,198). Elevations of α2,6 sialylation, specifically on integrins, have also been implicated in metastasis (216,217).

1.3.2 Glycomics

1.3.2.1 General Introduction into Glycomic Analysis

As glycosylation plays a pivotal in many diseases, numerous studies have attempted to characterize broad disease-associated alterations in the glycome. A complete survey of the entire glycome should determine the molecular composition of the glycan, the structure and branching pattern, and the anomeric linkages of each monosaccharide (218). These studies attempt to identify broad changes in glycosylation, which are ultimately representative of the partial or global proteome (219). While different approaches are commonly used to study the glycome, the general procedure typically involves [1] glycan release, [2] labeling, [3] enrichment/chromatography, and [4] detection and characterization.

1.3.2.2 Release of Glycans

Prior to analysis of the glycome, glycans must be released from the protein by enzymatic or chemical methods. Enzymatic release of glycans is primarily reserved for N-linked glycans due

to the availability of enzymes. Endoglycosidases are used for glycan release, while exoglycosidases can be used for structural analysis of released glycans. Endoglycosidases release the entire glycan, or a defined subset of the glycan, from the protein at the reducing end. For example, PNGaseF severs the bond between the asparagine amino acid and the glycan (220). PNGaseF is capable of releasing high-mannose, complex and hybrid glycans, with the only limitation being those with an α1,3-linked core fucose. However, the core α1,3 fucose linkage is not present in mammals, so the limitation is not a factor in the digestion of human samples. Endoglycosidase H (Endo H), an enzyme that cleaves between the two core GlcNAc residues on an N-glycan, can be used to digest high-mannose and hybrid glycans (220,221). Non-fucosylated glycans released by PNGaseF and EndoH will vary by the mass of a GlcNAc residue. Other endoglycosidases (i.e. EndoM, EndoF1, EndoF2, EndoF3) release classes of N-glycans depending on fucosylation and amount of branching (221). There are few commercially available enzymes that can release O-linked glycans, and none can universally release all classes of O-linked glycans (220). For instance, endo-α-N-acetylgalactosaminiadase, is only capable of releasing core 1 O-glycans (222). Thus, studies that profile the O-glycome rely on chemical methods for glycan release, primarily β-elimination and hydrazinolysis (223). The original protocol for reductive β-elimination utilizes basic sodium hydroxide conditions to cleave O-glycans under moderate temperatures, followed by the reduction of the terminal monosaccharide to an alditol (224). A faster release of O-glycans can be achieved by using dimethylamine instead of sodium hydroxide (225). Reductive β-elimination is not amenable to certain glycan detection strategies, as glycans cannot be labeled with a fluorophore or chromophore, as will be discussed in the glycan derivatization section. Approaches to release O-glycans by nonreductive β-elimination have been explored, but are subject to glycan degradation at the reducing end, referred to as "peeling" (226–228). Hydrazinolysis represents another nonreducing glycan release

strategy that can be applied to both N-and O-glycans (223). However, selectivity for O-glycan release can be achieved under lower temperatures (229).

1.3.2.3 Glycan Derivatization

Derivatization of native glycans is almost universally performed prior to glycan analysis. While native glycans can be profiled by modern instrumental methods, derivatization functions to [1] allow glycans be detected by alternative approaches, [2] stabilize monosaccharide residues that are labile under native conditions, [3] improve detection sensitivity, and/or [4] provide more informative structural information following tandem mass spectrometry (230). The most common N-glycans derivatization strategies can be classified into the broad categories of reductive amination, Michael addition, hydrazide labeling, permethylation, and the selective modification of terminal monosaccharides. In reductive amination, the reducing end of the glycan is labeled by a condensation reaction with a primary amine, resulting in a Schiff base that is then reduced to yield a secondary amine (230). Reductive amination is commonly used in glycomic analysis is to incorporate fluorescent tags on the reducing end of glycans, thereby allowing detection by a fluorescence detector after chromatographic separation. This strategy is compatible with a wide variety of labels, including 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA) and 2-aminopyridine (2-PA). Efforts to monitor the process of 2-AB and 2-AA labeling demonstrate a high labeling efficiency and negligible loss of sialic acid (231). However, glycans released by reductive β-elimination are incompatible with derivatization by reductive amination as the glycan is already reduced during glycan release. Michael addition labels the reducing end of the glycan, but under alkaline conditions reducing the loss of sialic acids. 1-Phenyl-3-methyl-5-pyrazolone (PMP) is the most common labeling reagent added through Michael addition, but no fluorescent labels are known (230). In an effort to improve throughput, a one-pot method to release and label O-glycans by β-elimination has been developed (232). Hydrazide labeling at the reducing end of the glycan results in a permanent positive charge on the glycan. Other labeling strategies are

directed at the non-reducing end or terminal monosaccharides. Permethylation replaces hydroxyl, amine and carboxyl groups with methyl groups, generating a more hydrophobic glycan, which can dramatically improve downstream chromatography and analysis (233). During the first step of this procedure, alcohol groups on the glycan are converted into alcoholate ions by sodium hydroxide in DMSO. A methyl group from methyl iodide is then substituted onto the alcoholate ions (220). In one adaptation of the method, the processes of O-glycan release and permethylation are combined (234). Other derivatization approaches selective modify the monosaccharides that complicate downstream analysis, specifically sialic acids. P-toluidine can incorporate onto the sialic acid residues on glycoprotein and the derivatized glycans can then be released with PNGaseF (235). Other approaches offer the ability to differentially modify sialic acids based off of their linkages. Methyl esterification results in the lactonization of α2,3-linked sialic acids and the formation of methyl esters on α2,6-linked sialic acids (219,236). A more recent study examined methanol, ethanol, 2-propanol, and 1-butanol as alkyl donors. Both methanol and ethanol performed well in initial testing, but ethanol [ethyl esterification] was concluded to exhibit more selectivity between α2,3- and α2,6-linked sialic acids (237). The modification of sialic acids enhances the stability of sialic acid residues, while enabling detection in positive ion mode.

1.3.2.4 Glycan Enrichment/Chromatography

The glycan enrichment/chromatography step serves two primary roles in the analysis of the glycome. The first function is the removal of salts and labeling reagents, which are often present in significant excess over the glycans. Solid phase extraction (SPE) approaches reduce the amount of labeling reagents present after glycan derivatization, especially for reductive amination or hydrazide labeling approaches (230). Liquid-liquid extraction is used to remove additives from permethylated glycans (238). The second function of chromatographic separation of native or derivatized glycans is to separate the glycans prior to detection, thereby reducing suppression of

lower abundance glycans or enabling structural isomers to be distinguished (239,240). Efficient chromatographic separation often relies upon on-line chromatography interfaced with the method of detection. Native glycans are hydrophilic and are not efficiently retained on reverse phase (RP) columns, but derivatization often increases the hydrophobicity of glycans and enables the use of RP chromatography for separation (239,241,242). In addition to RPLC, the hydrophilic properties of native glycans and some derivatized glycans encourages the use of hydrophilic interaction chromatography (HILIC), where there is a direct correlation between retention time and the size and polarity of the glycan (241,243). Ultimately, chromatography steps that employ RPLC are advantageous as labs that are currently equipped for bottom-up proteomics can process glycans with minimal alterations to the existing workflows already in place. Porous graphitized carbon (PGC) resin can enrich glycans in both on-line and SPE workflows, with enrichment following the same principles as HILIC enrichment (239,240,244). Thin layer chromatography (TLC) and gel electrophoresis (GE) can be used in place of LC to chromatographically resolve glycans (245). The emphasis placed on chromatographic separation of glycans has resulted in novel modalities developed specifically to study glycosylation, such as the GlycanPac AXH-1 analytical column, which can separate glycans based on charge, size and polarity (243). While the interaction is similar to HILIC chromatography, the charge state of the glycan does not play a role in HILIC enrichment.

1.3.2.5 Glycan Detection and Characterization

Traditional glycomic approaches rely upon chromatographic retention in comparison to reference glycan libraries from standard samples (239). For this approach, glycans are labeled with a fluorophore or chromophore, which not only stabilizes the glycan but also serves to enable detection by fluorescent or UV detectors following chromatography (223). In addition to retention time, glycan structures can be validated by sequential glycosidase digestions (219,246). After exoglycosidase digestion, a change in the retention time denotes the presence of the enzyme

substrate that is targeted by the exoglycosidase. While useful, this approach is limited by the number of commercially available enzymes. Complete structural assignment may be unobtainable as enzymes, such as those to specifically digest bisecting GlcNAc residues, are unavailable (246).

More recently, glycomic analysis has benefitted from advances in the field of mass spectrometry, as mass spectrometry is more sensitive and consumes less sample. Importantly, studies have shown that the signal derived from neutral and permethylated glycans increases with glycan abundance (233,247). MALDI-MS is a fast and convenient way to profile the glycome. Both neutral and acidic native glycans can be profiled by MALDI but the presence of both types in a single sample complicates analysis (231). In positive ion mode, neutral glycans are predominantly detected as $[M + Na]^+$ ions, but $[M + H]^+$ and $[M + K]^+$ are also observed (220). However, MALDI-MS in positive ion mode is the not ideal for detecting sialylated glycans, as sialic acid residues are often labile and released by in-source decay due to the energy of the laser and vacuum conditions in the ion source (248,249). The use of ion sources with higher-pressure conditions reduces the dissociation of sialic acids. Negative ion mode is primarily used for the detection of sialylated glycans, which are detected as $[M - H]^-$ ions, as opposed to $[M + Na]^+$ and $[M - nH + (n + 1Na)]^+$ ions in positive ion mode (247). Sialylated glycans profiled in negative ion mode do not suffer from sialic acid loss to the extent observed in positive ion mode. The choice of MALDI matrix has a pronounced effect on the signal observed for the detection of native glycans. Both DHB and CHCA can be used to survey glycans in positive ion mode, with CHCA promoting more sialic acid dissociation (250). Furthermore, the combination of matrices has been demonstrated to improve glycan analysis over individual matrices. DHB/CHCA and DHB/SA combinations were shown to effectively increase ionization efficiency of glycans and improve shot-to-shot reproducibility (251). Additionally, DHB/SA enables detection of both glycans and deglycosylated proteins in a single experiment. In addition to MALDI, ESI, typically preceded by chromatographic separation, is also widely used for glycomic analysis. However, ESI generates a

weaker signal compared to peptides and proteins, in part due to limited surface activity in ESI droplets caused by the hydrophilicity of the glycans (230,233). Using ESI, quantitation of N-glycans can be achieved using multiple-reaction monitoring (MRM) in the mass spectrometer (252).

It is often beneficial to derivatize glycans prior to analysis by mass spectrometry. The increased hydrophobicity of glycans as a result of permethylation increases the ionization efficiency and results in higher signal and increased stability. As permethylation functions by substituting methyl groups for hydrogen groups on hydroxyl, amine, and carboxyl groups, permethylated glycans do not differ by a consistent mass difference compared to the native glycans. Instead, a larger glycan will have a relatively larger shift in molecular weight, due to increased substitutions sites on the glycan. When derivatization only occurs at the reducing end of the glycan, the resulting mass shift will not vary with the size of the glycan. An example of this is the 2-AB label, which is commonly used for fluorescence detection, but is also compatible with MALDI and ESI (231). The labeling of labile and acidic residues can also consolidate the polarity used for glycan detection. Labeling by permethylation or p-toluidine enables sialylated glycans to be detected in positive ion mode (230,235).

The initial assignment of the glycans present in a sample is often done by mass accuracy or retention time, but an additional structural validation is sometimes required (253). As previously mentioned, sequential exoglycosidase digestion represents a traditional approach to study monosaccharide linkages. This technique is often employed prior to chromatography, but can be implemented prior to mass spectrometry as well (219,246). Given the limitations in throughput and enzyme availability, mechanisms to structurally differentiate glycan isomers by mass spectrometer have been investigated. Fragmentation in a mass spectrometer represents a less time consuming and easier technical approach to characterizing the composition and linkages of a glycan. Initial fragmentation techniques relied on spontaneous post-source decay, taking

advantage of the poor stability of glycans, albeit usually at a poor resolution (218,247). While useful for verifying the glycan composition, this approach was unable to produce information relevant for defining the structure. More recently, CID is more commonly utilized. It is important to understand information produced by specific ions when evaluating the efficiency of fragmentation for structural elucidation of glycans. The original nomenclature for glycan product ions, established by Domon & Costello, defined two series of fragment ions; a, b and c, and x, y, and z (254). The first series of ions (a, b, and c) contain the non-reducing terminus, while the second series (x, y, and z) contain the reducing terminus (Figure 5). The b, c, y and z ions are generated from fragmentation across the glycosidic bonds, while a and x ions are generated from cross-ring cleavages (254). Defining anomeric linkages by fragmentation requires the generation of a and x ions. Unfortunately, when selected for fragmentation by low energy CID in positive ion mode, $[M + H]^+$ and $[M + Na]^+$ ions primarily fragment across the glycosidic bonds, with the sodium adduct providing more linkage-informative cross-ring cleavages. Ramping up the fragmentation energies promotes more cross-ring cleavages (255). While not ideal, fragmentation across glycosidic bonds can sometimes provide structural information based off of characteristic diagnostic ions, such as those generated from the presence of a bisecting GlcNAc residue or core fucose (247,255). Efforts to improve the fragmentation efficiency and formation of cross-ring cleavages have focused on the using sample additives, different MALDI matrices, and glycan derivatization. The addition of alkali metal additives into the glycan sample enhance fragmentation efficiencies (Li>Na>K>Rb>Cs) and activation energies (256). For MALDI-MS analysis, the choice of matrix significantly impacts the glycan fragmentation pattern. CHCA is characterized as a “hot” matrix and primarily promotes cleavages at the glycosidic bonds, predominantly forming b and y type ions (218). In contrast, DHB provides a more informative fragmentation pattern, promoting glycosidic and cross-ring cleavages. The glycan structural information can be further enhanced by glycan derivatization. CID in negative ion mode is quite

productive in structurally defining glycans. Fragmentation of glycans in negative ion mode promotes more cross-ring cleavage and diagnostic ions that denote branching patterns, the presence of bisecting GlcNAc residues, and the location of fucose residues. As with positive ion mode, fragmentation can be enhanced by the addition of anions, such as chloride, nitrate, and phosphate salts (246).

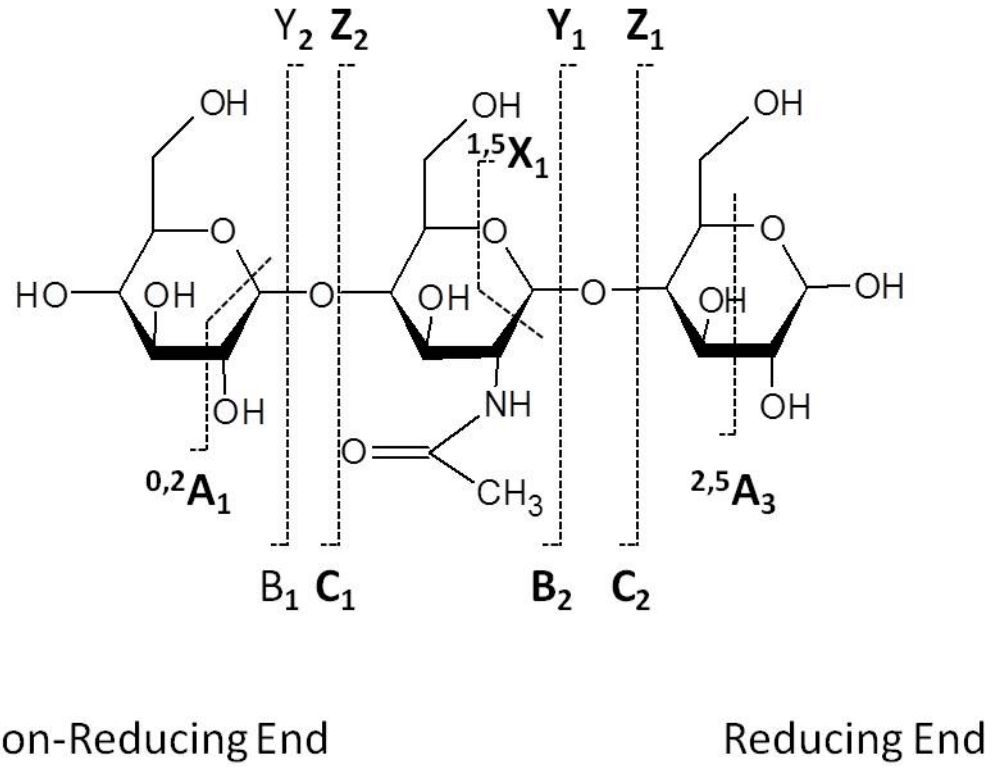


Figure 5. Glycan Fragmentation Series. Fragmentation of glycans results in different series that either contain the reducing end (x, y, and z ions) or do not contain the reducing end (a, b, and c ions). The types of ions generated often depend on the fragmentation strategy used.

1.3.3 Glycoproteomics

1.3.3.1 General Introduction into Glycoprotein Analysis

Given the dynamic role of glycosylation in biological and pathological functions, glycoproteomics has become extremely relevant in the initial phase of biomarker discovery

research. Protein glycosylation is highly specific, both in regards to the site of glycosylation and the oligosaccharide structures. However, disease related changes in glycoproteins can manifest as one or multiple of the following characteristics; [1] alterations in the site or frequency of glycosylation, [2] overall changes in protein abundance, or [3] changes in structure of the oligosaccharides at a specific site. Thus, an uncoupled analysis of the glycome and proteome results in a significant void of information by not linking the proteins to their attached glycans. General glycomic analysis results in a loss of information about the modified protein and specific sites of glycosylation, while proteomic analysis ignores the site and structure of oligosaccharide modifications. Glycoproteomics aims to rectify these limitations by characterizing and quantifying the protein and glycan attachments for each glycoprotein in the sample. Glycoproteomic research is typically directed at secreted proteins detected in bodily fluids (i.e. saliva, cysts, serum, secretions from cells) or membrane proteins (i.e. tissue sections or cell membrane fractions). Generally, secreted glycoproteins represent ideal biomarkers for the non-invasive detection of diseases, while membrane glycoproteins provide information on the cellular mechanisms that drive the disease (257).

1.3.3.2 Glycoprotein Enrichment

1.3.3.2.1 General Enrichment Overview

While recent analytical advances have improved the analysis of the glycoproteome, significant limitations in mass spectrometry-based glycoproteomics persist due to the inherent complexity of glycosylation. This complexity is attributed to poor digestion efficiency, the presence of non-glycosylated peptides, glycan heterogeneity, and the physiochemical properties of the glycan modification. A majority of glycoproteomic research relies upon bottom-up proteomic analysis of glycopeptides or previously glycosylated peptides, due to the limitations of top-down mass spectrometry (129,258). Trypsin is the most common enzyme used for glycoprotein digestion, but often produces prohibitively large glycopeptides due to limited

digestion sites and missed cleavages when the glycan is in close proximity to the cleavage site (259,260). Thus, the use of other enzymes with complementary cleavage sites can improve the number of glycoproteins detected (261). While it is estimated that >50% of human proteins are glycosylated, only 2-5% of peptides are glycopeptides, thus there is a significant need to enrich glycopeptides from the total peptide mixture (262). Another complication for glycoproteomics is the issue of protein dynamic range. Serum proteomics for low abundance biomarkers often necessitates multi-dimensional separation strategies to overcome the large dynamic range of serum proteins, which is approximately $10^{\wedge 12}$ (263). It is estimated that 22 of the most abundant plasma proteins constitute 99% of the total protein mass, including some (i.e. albumin) which are non-glycosylated (150,221,264). As these proteins provide little clinically relevant information, depletion often improves the total number of identified proteins (63). Even though the dynamic range is not as robust in tissue sections, enrichment strategies are still required (258). Macroheterogeneity and microheterogeneity further enhance the complexity of glycoprotein analysis (221,265). In conjunction with dynamic range issues, micro- and macroheterogeneity result in a low abundance of identical glycopeptides, especially compared to peptides with no posttranslational modifications. Finally, the hydrophilicity of glycans reduces the ionization efficiency of the peptides when analyzed by mass spectrometry. Taken together, the low abundance, complexity and low ionization efficiency of glycopeptides necessitate glycopeptide enrichment prior to mass spectrometry. While enrichment can be performed at the protein level, enrichment at the peptide level is beneficial as some glycans are masked by the protein tertiary structure (257). Enrichment of glycoproteins can be performed at the protein and peptide level and can be classified into three broad categories; affinity enrichment, covalent interactions, and physiochemical enrichment methods (266).

1.3.3.2.2 Affinity Enrichment

Affinity enrichment approaches rely on lectins or antibodies that recognize specific glycan epitopes or monosaccharide linkages. Under both strategies, glycoproteins that interact with the lectin or antibody are retained, while other molecules are removed by washing. Lectins are carbohydrate-binding proteins that recognize specific oligosaccharide epitopes with various degrees of specificity. Lectin enrichment offers the capability to do both broad, or targeted glycoprotein enrichment, depending on which lectins are used. Broad specificity lectins, such as concanavalin A (ConA) or wheat germ agglutinin (WGA) display a broad specificity for common glycoforms. Other lectins interact with specific monosaccharides at defined linkages, such as sambucus nigra lectin (SNA), which interacts preferentially with α2,6-linked sialic acids (267). Even though not all anomeric linkages can be differentiated using lectins, the combination of lectins and sequential exoglycosidase digestion can provide relevant structural information based off of changes in glycoconjugate binding (268). A general redundancy exists in glycoprotein/ glycopeptide identification when comparing enrichment with different lectins, as glycoforms interact with multiple lectins (269). After enrichment, bound glycoconjugates are eluted by adding sugars with high affinity for the lectin to outcompete glycoconjugate binding. This reversible binding mechanism enables downstream analysis of intact glycoconjugates (270). The scope of glycoprotein analysis can be broadened by using multiple lectins simultaneously (MLAC) or sequentially (SLAC) to enrich glycoconjugates with broad glycosylation motifs (271). These approaches are used to provide a more global glycoconjugate enrichment or to enrich structures the do not interact with a single lectin. Total enrichment efficiency can be improved by incorporating nonionic detergents (221). While some glycoprotein enrichment approaches require the removal of abundant, non-glycosylated proteins (i.e. albumin), a comparison of lectin enrichment before and after depletion demonstrated little utility (269). Several limitations of lectin enrichment include variation due to poor standardization of lectins

from different commercial sources, poor reproducibility across large cohorts of samples, and low binding efficiency (269). When applied to global glycoproteomics, lectin enrichment may result in experimental bias due to variation in lectin-glycan binding affinity across glycoforms. In addition to glycoconjugate enrichment, the use of lectins in a microarray format can identify which lectins will be useful in an experimental study. Using this technique, proteins are fluorescently labeled and allowed to interact with immobilized lectin spots. Lectins that display altered binding between two conditions, such as tumor and non-tumor, represent promising targets to enrich aberrantly glycosylated proteins (221). This is especially useful as over 160 lectins have been identified and at least 60 are sold commercially (267). Commercially available lectins are available in pure, resin-bound, biotinylated and fluorescently-labeled forms (269).

Alternatively, antibodies targeted against glycan antigens can be used for glycoprotein enrichment. However, this approach is limited as glycans are poor antigens, resulting in few antiglycan antibodies with high affinity and specificity (272). Antiglycan antibodies have been utilized to enrich glycoproteins with O-GlcNAc, O-GalNAc, SLeX, CA19-9, and polysialic acid epitopes (221,273). However, the ability to generate antibodies requires the presence of multiple monosaccharides or the combination of a monosaccharide with a peptide. This has limited the success of generating antibodies specific to truncated glycoforms (274,275). Most antiglycan antibodies recognize terminal residues, but the underlying glycoform and class of glycan can play a role in binding affinity (150).

1.3.3.2.3 Covalent Interactions

Glycoconjugate enrichment strategies that rely on covalent bond formation improve the specificity of glycoconjugate enrichment by increasing the efficiency of non-glycoconjugate removal (276). These approaches enrich glycoconjugates by forming covalent bonds between the glycan moiety and a functionalized solid support. In the most established method, hydrazide chemistry, protein-bound carbohydrate *cis*-diol groups are oxidized to aldehydes, which then

react with the hydrazide resin to form hydrazone linkages. In the initial protocol, non-glycosylated proteins were removed by stringent washing, and bound glycoproteins were digested with trypsin to release non-glycosylated peptides from the glycoproteins. The bound glycopeptides were then removed by incubation with PNGaseF and analyzed by LC-MS/MS (277). More recently, this approach has been adapted for enrichment at the peptide level, which reduces sample complexity and exposes concealed N-glycopeptides to the hydrazide resin (262,278). Interestingly, there is poor overlap between protein identification when hydrazide is performed at the peptide level compared to the protein level (279). At the peptide level, the overall specificity of hydrazide-mediated enrichment of glycopeptides is improved (221,262). On the other hand, enrichment at the glycoprotein level is associated with a larger number of glycoprotein identifications (63,279). To date, hydrazide enrichment has been broadly adapted for glycoconjugate enrichment of cell lysates, fluids, and tissues (63,258,261,262,279–283). PNGaseF is used to release glycopeptides from the resin, resulting in asparagine deamidation, which can be used to identify the site of glycosylation. Misidentification caused by chemical deamidation, or deamidation not attributed to PNGaseF, can be reduced by performing the digestion in ^{18}O water, which results in a mass shift of 3Da (129,284,285). Furthermore, the use of higher resolution mass spectrometers and optimized search parameters can increase glycopeptide identification and reduce misidentification to around 1%. Unlike lectin enrichment, traditional hydrazide enrichment does not exhibit preference towards specific glycoforms, making it an ideal approach for global glycoprotein analysis. In one adaptation of the standard protocol, the use of mild oxidizing conditions can specifically enrich glycopeptides that contain a sialic acid (286). In this approach, acid hydrolysis is used to release the glycopeptide by cleaving the glycosidic bond between the sialic acid and the rest of the glycoform, enabling intact glycopeptide analysis (276). Given the importance of hydrazide enrichment in glycoprotein research, efforts to improve enrichment by altering wash conditions, enzymatic digestion, pre-

enrichment depletion, and membrane protein solubility have been investigated (63,261). Of note, the depletion of common proteins, and the use of detergents to solubilize membrane proteins, improved glycoprotein identifications. Incorporation of isotopically labeled compounds enables quantitative glycoprotein analysis (277). The primary limitation of hydrazide enrichment is that glycopeptides are released by PNGaseF, resulting in a complete loss of information about the structure of glycans on specific peptides. However, the deglycosylated peptides exhibit more efficient retention to RP columns and improved sensitivity in mass spectrometers, making glycoprotein identification easier than intact glycopeptides.

Boronic acid chemistry is another glycoprotein enrichment strategy that utilizes covalent binding (270). This enrichment strategy has been used for the enrichment of carbohydrates, nucleosides, glycolipids, RNA and glycoproteins. In boronic acid chemistry, basic conditions result in covalent ester bonds formation between the boronic acid and *cis*-diols on the glycans. These bonds are reversibly hydrolyzed under neutral or acidic conditions, rendering intact glycopeptide analysis possible (221,287). Like general hydrazide enrichment strategies, boronic acid enrichment is unbiased in terms of enriching specific glycoforms.

Glycoproteins can be enriched using the concept of "bioorthagonal click chemistry", a term coined by Carolyn Bertozzi, which refers to the incorporation of non-native molecules into biological systems without the impairment of biological processes (288). Importantly, the reaction involves the incorporation of a monosaccharide analog with an azide functional group, and a subsequent interaction with an exogenous probe that reacts with the analog through click chemistry, allowing for detection and enrichment of the target (289). The utility of the azide functional group lies in its small size, biological stability, and absence in native biological systems. These traits allow the azide-labeled monosaccharide analogs to incorporate into glycoconjugates using the cell's natural biosynthetic machinery, without modifying the underlying protein structure or function (289,290). Acetylation of the sugar analog increases the

lipophilicity and allows the monosaccharide to passively diffuse through the plasma membrane, which can then be deacetylated by nonspecific esterases (291). Once incorporated, the sugar analogue can be assessed by reactions that involve azide-alkyne cycloaddition or Staudinger ligation. Global changes in glycosylation can be accessed by a variety of analytical approaches, including flow cytometry, confocal microscopy, and gel electrophoresis (292,293). For mass spectrometry-based proteomic applications, glycoproteins can be enriched using alkyne resin, which reacts with the azide functional group forming a 1,4 disubstituted 1,2,3-triazole linkage (294). The use of alkyne-agarose resin, as opposed to biotin-alkyne reduces the number of enrichment steps and results in higher glycoprotein specificity (294,295). Metabolic labeling of glycans was initially optimized for *in vitro* labeling glycans, but has more recently been applied to *in vivo* models (296–300). Multiple sugar analogs have been used to enrich glycoproteins with various monosaccharides compositions. For instance, tetraacetylated N-azidoacetyl-D-mannosamine (ManNAz) incorporates into glycoproteins in place of sialic acids. The incorporation of ManNAz onto sialic acid containing glycoproteins was validated by Western blot analysis before and after sialidase treatment (301). Tetraacetylated N-azidoacetylglucosamine (GlcNAz) can be utilized to survey O-GlcNAc-modified proteins (192,298). While the efficiency of incorporation is low, treatment with O-GlcNAcase inhibitors has been reported to improve O-GlcNAz incorporation. Tetraacetylated N-azidoactylgalactosamine (GalNAz) was originally reported to incorporate in place of GalNAc monosaccharides, but can also incorporate in place of O-GlcNAc modifications following the conversion of UDP-GalNAz to UDP-GlcNAz by UDP-galactose 4'epimerase (GALE) (192,297). Interestingly, GalNAz is more efficient at labeling sites of O-GlcNAc modifications than GlcNAz, which is attributed to the rate-limiting UDP-GlcNAc pyrophosphorylase step in the GlcNAc salvage pathway (192). Peracetylated 6-azidofucose (6AzFuc) is converted intracellularly to GDP-FucAz prior to incorporating in place of fucose (302). Conversely, treatment directly with GDP-FucAz improves the incorporation efficiency by

bypassing the fucose salvage pathway (289). For all analogs, the incorporation efficiency varies amongst different cell lines (290). During the click chemistry enrichment process, covalently bound glycoproteins are separated from non-glycosylated proteins following extensive washing steps. After the removal of non-glycosylated proteins, glycoproteins are digested with trypsin enabling protein identification by LC-MS/MS. Glycopeptides remain bound to the click resin, but can be assessed following PNGaseF digestion. Unfortunately, given the nature of peptide release, all information regarding the exact structure of the glycoform is lost.

1.3.3.2.4 Physiochemical Enrichment

Glycans substantially alter the physiochemical properties of the underlying peptide. Unlike non-glycosylated peptides, glycopeptides are poorly retained on the C18 stationary phase, due to their increased hydrophilicity (221). This causes most glycopeptides to elute early during chromatography, with little resolution of unique glycopeptides. Even glycopeptides with different glycoforms are not adequately resolved on RP columns, with the exception being those that differ in their levels of sialylation (303). However, RP-HPLC-MS remains integral to glycoproteomic workflows involving the analysis of deglycosylated peptides. Alternatively, RP-LC can be used in conjunction with other chromatographic approaches that are able to resolve glycopeptides (303–305). These other chromatographic approaches are better suited to enrich and resolve glycopeptides.

Hydrophilic interaction chromatography (HILIC) is used to enrich all glycopeptides, irrespective of the glycan structure. HILIC enrichment is based off of the increased peptide hydrophilicity due to the bound glycan. Glycopeptides are retained on the hydrophilic stationary phase, and are eluted with an increasingly aqueous buffer (267). While mechanistically similar to normal phase (NP) chromatography, analyte retention on HILIC columns is based on the association of the analyte with a water layer surrounding the polar stationary phase, instead of the stationary phase itself (306). Multiple materials are amenable to HILIC enrichment of

glycopeptides, including zwitterionic (ZIC), amide-based (TSKgel Amide-80), and carbohydrate-based materials (cotton wool, sepharose) (307). ZIC-HILIC enrichment involves polar and ionic interactions, while cotton HILIC materials utilize hydrogen bonding (308). In a comparison between HILIC materials, ZIC-HILIC provides higher selectivity and efficiency in glycopeptide enrichment. The use of both materials in tandem further improves glycopeptide selectivity (303,308). Unlike RP-LC, HILIC-LC is able to resolve glycopeptides that vary only in the composition of neutral monosaccharide residues (303). When applied in a SPE format, HILIC material is able to enrich glycopeptides and remove detergents, prior to downstream LC-MS applications (309). A drawback of HILIC enrichment is the relatively poor enrichment specificity compared to other enrichment strategies. Hydrophilic non-glycosylated peptides are often retained during HILIC enrichment, while more hydrophobic glycosylated peptides are lost (310). To improve the specificity of glycopeptide enrichment, TFA can be added as an ion-pairing reagent, which acts to decrease the hydrophilicity of non-glycosylated peptides, while not affecting the hydroxyl groups of the glycan on glycopeptides. The net effect of using ion-pairing agents is to further separate the hydrophilicity of glycopeptides from non-glycosylated peptides (306,308).

Physiochemical enrichment strategies can be applied to enrich selected glycopeptides. Glycopeptides with sialic acids can be enriched with TiO_2 through the interaction of negatively charged carboxylic acid and hydroxyl groups of sialic acids with Ti^{4+} (311). TiO_2 also enriches phosphopeptides, so the specificity of glycopeptide enrichment is improved by pretreating with phosphatases (267,311). Larsen et al. compared glycopeptide enrichment pre- and post-neuraminidase treatment to demonstrate the importance of sialic acid in TiO_2 enrichment (311).

New approaches to enrich glycopeptides are continually being evaluated. Takakura et al. demonstrated the ability to selectively precipitate glycopeptides over non-glycosylated peptides in acetone (312). Preliminary unpublished evidence demonstrates the utility of strong anion

exchange (SAX) for glycopeptide enrichment, with initial results showing improved enrichment compared to HILIC. Finally, Alvarez-Manilla et al. used an *in-silico* analysis of all human peptides to propose size exclusion chromatography as an enrichment strategy for glycopeptides, based on the added molecular weight imparted on the peptide by the glycan. In their analysis of digested peptides, 90% of peptides are smaller than 2000 Da, while nearly all glycopeptides are larger than that threshold (313).

1.3.3.2.5 Selection of Enrichment Strategies and Combinatorial Approaches

The determination of an ideal enrichment strategy depends on the ultimate goal of the study. Most published studies of the glycoproteome focus primarily on the N-glycoproteome as opposed to the O-glycoproteome due to [1] reduced protease shielding from fewer sites of glycosylation, [2] predictable N-glycan consensus sequences, [3] a common core structure of N-glycans, [4] and a larger size of N-glycans resulting in more pronounced physiochemical changes (257). For cancer studies, the appropriate enrichment strategy is determined by preceding data implicating differential glycosylation; including differential lectin staining, altered glycosylation-related gene expressions, or changes observed at the glycan level. This preexisting data may dictate whether it is beneficial to enrich specific glycoforms as opposed to all glycoproteins. Efforts directed at identifying both the glycoform and underlying peptide backbone typically gravitate towards the use of lectins or physiochemical enrichment strategies, due to the reversible mechanism of glycoconjugate release. However, these approaches do not provide the specificity of covalent-mediated glycopeptides enrichment. In a comparison of enrichment methods in the cell secretome, Li et al. reported a higher selectivity of glycopeptides with hydrazide enrichment (80% specificity) compared to ZIC-HILIC enrichment (30% specificity) (314).

Combinations of enrichment strategies can enhance the specificity of glycoconjugate enrichment and the total number of identified glycoproteins. Combinations can use similar enrichment approaches, such as the combination of multiple lectins, or different enrichment

strategies. There is often little overlap in the glycopeptides and glycoproteins identified using different enrichment strategies, resulting in more glycoprotein identifications (282,314). For instance, a comparison of hydrazide and lectin enrichment approaches for blood serum revealed only a 42% overlap in common glycoproteins (282). The differential enrichment can be explained in part by the specificity of lectins and antibodies for a subset of glycan structures compared to the non-selective hydrazide approach.

1.3.3.3 Mass Spectrometry Strategies for Glycoprotein Identification

1.3.3.3.1 Ionization Strategies

Following glycoconjugate enrichment, the selection of the proper ionization source and fragmentation parameters can dramatically improve glycoprotein identification and reduce analytical complexity (271). As previously referenced, glycosylated peptides are not ionized as efficiently as their deglycosylated counterparts. Multiple studies have compared the utility of ESI and MALDI for intact glycopeptides. MALDI-TOF instruments operating in linear mode result in broad peaks that contain information on the average carbohydrate content at the glycosylation site (249,315). The removal of glycans or further digestion of the underlying peptide is required to obtain more informative information (315). When performed in reflectron mode, MALDI-TOF is able to profile glycopeptides with isotopic resolution (249). Studies have shown that the overall signal strength of glycopeptides when assessed by MALDI-TOF correlates with the relative abundance, regardless of the peptide backbone (316). Unfortunately, in-source decay of the oligosaccharide frequently occurs during MALDI ionization, specifically at sialic acid and fucose residues, rendering the use of derivation approaches essential. The use of MALDI-FTICR instead of MALDI-TOF improves the glycopeptides resolution and aids in the retention of sialic acids (249). ESI is the more common ionization approach for intact glycopeptides analysis. As ESI is a more gentile ionization strategy than MALDI, fragmentation is less common (315). Additional advantages are that ESI ionization is more efficient than MALDI, can process glycopeptide

mixtures quickly following chromatographic separation, and can be used to trigger MS^n scans (317). One disadvantage for ESI is that the ionization results in multiply charged analytes, rendering the spectra more difficult to process manually. Irungu et al. found that the two ionization sources are quite complementary, with ESI-MS/MS generating the most complete information on the glycan moieties and MALDI-MS/MS providing high confidence peptide assignments (317).

1.3.3.3.2 Fragmentation for Glycopeptide Identification

The need to characterize both the glycan and underlying peptide backbone complicates mass spectrometry approaches for glycopeptides. CID fragmentation is insufficient to fragment both the glycan and peptide. Complete fragmentation of the glycan moiety is achieved at lower collision energies that leave the underlying peptide backbone largely intact (267,318,319). The peptide-glycan bond is also cleaved at lower collision energies diminishing information about the site of glycosylation (320). At higher collision energies, only peptide backbone b and y fragmentation series are observed. Thus, other fragmentation approaches have been investigated for intact glycopeptide analysis. Electron-activated dissociation methods, such as ETD and ECD, fragment the peptide backbone into c and z fragments, while retaining the glycan on the original site of glycosylation. Therefore, peptide sequencing can be performed using the c and z type fragmentation series, while the glycan is still bound to the glycosylation site (320). Conversely, HCD primarily fragments the glycan, generating oxonium ions specific to monosaccharides. Importantly, HCD in the C-trap and detection in the Orbitrap is not limited by the low mass cutoff associated with ion trap mass spectrometers, and is thus able to overcome the 1/3 cut-off rule to enable the detection of low molecular weight oxonium ions. While fragmentation of the underlying peptide is generally inefficient, HCD does generate a characteristic y1 ion that is representative of the peptide + GlcNAc (221,267). HCD fragmentation and subsequent detection

in the orbitrap, is also compatible with tandem mass tag (TMT) labeling for quantitative mass spectrometry analysis (319).

Due to the inadequacies of single ionization approaches in determining the site of glycosylation and the structure of both the glycan and peptide, multiple ionization approaches have been used in succession for the analysis of glycopeptides. One such strategy utilizes the Thermo Orbitrap Elite mass spectrometer to combine HCD and ETD in a product dependent manner (HCD-pd-ETD) (Figure 6). Using this approach, abundant precursor ions are selected for HCD fragmentation. HCD fragmentation of glycopeptides results in the generation of diagnostic oxonium ions, which are then detected by the mass spectrometer and trigger ETD fragmentation of the original precursor ion (320). Information gathered by the HCD and ETD fragmentation can be combined to provide relevant information on both the glycan and underlying peptide. This approach improves the duty cycle of the mass spectrometer in comparison to alternating HCD and ETD (HCD-alt-ETD), by initiating ETD fragmentation only on glycopeptides. In a comparison of HCD-pd-ETD to HCD-alt-ETD, Yin et al. demonstrated that HCD-pd-ETD was more useful in identifying low abundance glycopeptides, while HCD-alt-ETD predominantly identified high intensity glycopeptides (321). However, both approaches identified fewer total proteins than analysis of deglycosylated peptides. Additionally, only 16 out of 151 glycoproteins were identified by all three strategies (321). Improvements in analytical instrumentation continue to improve glycopeptides analysis. The Orbitrap Fusion outperforms the Orbitrap Elite by enabling data acquisition on the highest charged precursors, which have the most robust fragmentation by ETD, as opposed to the most intense ions. In contrast to HCD-pd-ETD, HCD-pd-CID can better elucidate the glycan structure (322). The new trihybrid Orbitrap configuration allows CID to also be incorporated into the HCD-pd-ETD workflow. This HCD-pd-CID/ETD fragmentation workflow reduces glycan assignment errors, provides an additional validation step, and can provide structural assignments based off of the CID fragmentation pattern (322).

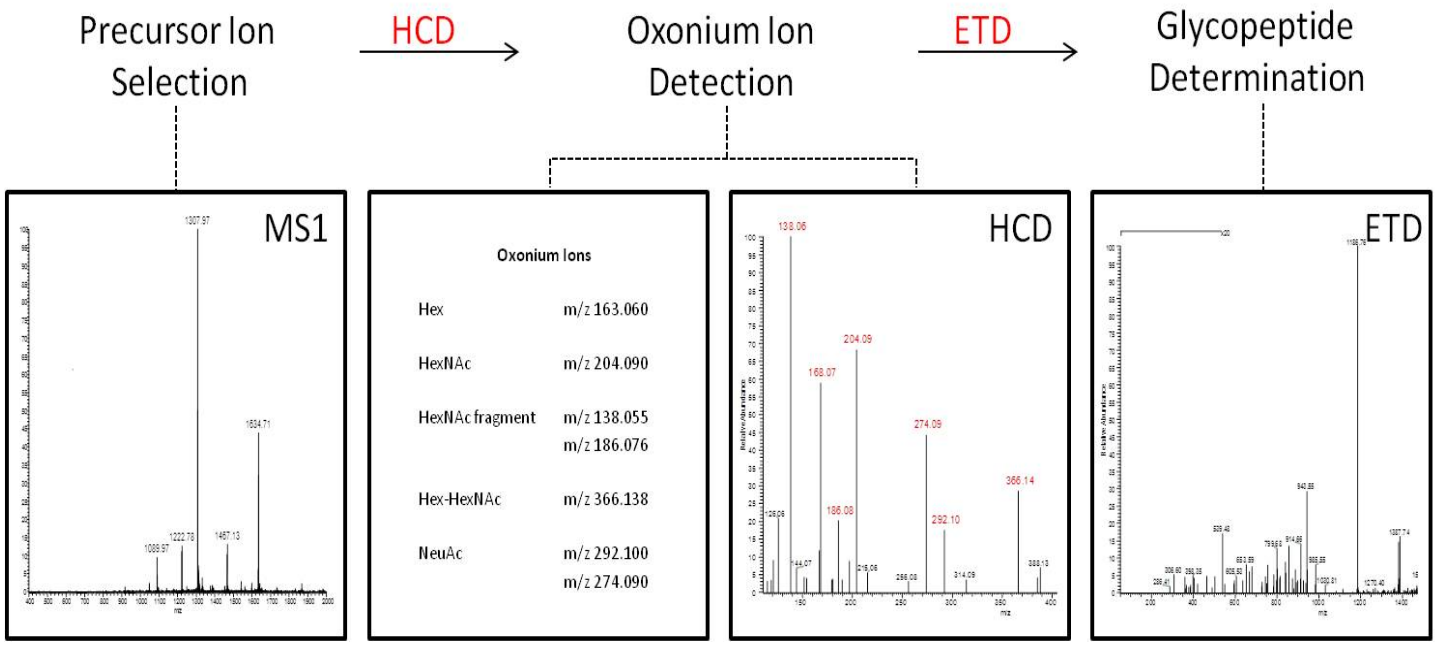


Figure 6. HCD-pd-ETD Workflow. Abundant precursor ions are selected for fragmentation by HCD, which generates diagnostic oxonium ions. The presence of oxonium ions triggers ETD on the original precursor ion, which is used for glycan, peptide, and protein characterization.

1.3.4 Glycoproteins and Glycans Antigens as Biomarkers

1.3.4.1 Glycoprotein Cancer Biomarkers

The detection of cancer at early stages of tumor progression is advantageous from a therapeutic treatment standpoint, resulting in improved treatment efforts and survival statistics (161). Disease markers, often referred to as biomarkers, can be utilized to indicate normal biological functions, disease processes, or response to therapy (323). The detection of one subset of biomarkers, tumor markers, which arise from either the tumor or the host, indicate the presence of a tumor (324). Ideal tumor markers should have very high performance in terms of sensitivity, specificity and predictive value, originate from specific organs instead of being ubiquitously expressed, and correlate with disease prognosis; including the tumor stage, volume, and likelihood of recurrence (325,326). For early stage cancer diagnosis, markers should ideally present prior to malignancy. Disease markers have been classified as proteins, glycoproteins, glycan antigens, hormones, hormone receptors, enzymes, genetic mutations or aberrations, and RNA (324,326). However, when looking at protein-based biomarkers, nearly all FDA approved biomarkers for cancer are glycoproteins or individual glycan antigens (160). Biomarkers experiments directed at identifying glycoconjugates are popular due to [1] the success for these molecules as biomarkers, [2] the prevalence of these molecules in biological fluids, and [3] the role that glycosylation plays in tumor progression. Clinically, these markers can be detected in a variety of complex biological matrices, including serum, plasma, urine and tissue. While research has been directed at identifying new tumor markers, very few prospective markers progress to clinical application. The poor transition from biomarker discovery to clinical assay can be explained by a variety of factors, including technical limitations, sample complexity, and experimental design. Techniques utilized to identify disease markers, including mass spectrometry based "omics" strategies, are not easily implemented in clinical labs (324).

Additionally, biomarker assays directed at individual proteins suffer from the broad dynamic range of protein expression, with biomarkers often expressed at very low levels compared to abundant serum proteins (325). Changes in glycosylation resulting from malignancy are also difficult to detect and obscured by more abundant glycoforms (160,327). Among other reasons, prospective biomarkers rarely transition to clinical biomarkers because research scientists do not have the knowledge or experience needed to navigate the regulatory requirements necessary to advance the biomarker to a clinical assay. To adjust for these limitations, the National Research Council Early Detection Research Network (EDRN) has laid out a series of phases for biomarker development (324). Narimatsu demonstrates the workflow by taking a small number of potential disease markers identified in tissue, assessing their presence and serum, and then developing a simplified assay to assess the most promising biomarker on a large number of serum samples in a high-throughput fashion (327). While it is ideal for biomarkers to work with non-invasive biological fluids, tissues are frequently used to identify initial disease marker candidates. In comparison to serum, tissues have a reduced protein dynamic range and a direct relationship with the origin of the biomarker. As described above, these initial targets can be validated in fluids (328). Clinical diagnosis may require the use of tissue/biopsy cores, but biofluid analysis is preferable, as the process of obtaining the biopsy is invasive and risky.

While biomarkers can be difficult to identify in initial biomarker discovery approaches, targeting specific molecules with known identities is a more effective strategy. Mass spectrometry represents a powerful approach for initial serum/tissue biomarker investigations, but is often too complicated, expensive, and time-consuming to be utilized in clinical labs. The most common assays used in the clinic to measure biomarkers are immunohistochemistry (IHC) for tissue sections, or immunoassays for biological fluid samples. Assays to measure glycoprotein biomarkers in fluids can be compartmentalized into 3 categories; [1] the use of monoclonal antibodies directed against the protein, [2] the use of antibodies or lectins to measure the level of

an antigen/glycan motif, and [3] a combination of the previous approaches to quantify both the protein and associated glycoforms (325). An example of each category of biomarker is described in the text below by prostate-specific antigen (PSA), CA19-9, and alpha fetoprotein (AFP), respectively. The use of monoclonal antibodies to determine protein expression primarily relies on enzyme-linked immunosorbant assays (ELISA) (326). Similarly, enzyme-linked lectin assays (ELLA) or monoclonal antibodies directed against glycan antigens are used to measure total glycan antigen levels. The combination of protein and glycoform quantification can be achieved by either capturing all glycoproteins with immobilized lectins and then assessing the abundance of a specific protein with antibodies, or using an antibody to capture a specific protein and then detecting the abundance of the glycoform with lectins (325,327).

1.3.4.2 Selected Glycoprotein Biomarkers

1.3.4.2.1 Prostate-Specific Antigen (PSA)

PSA is a protein secreted from the prostate that is detected in the blood at small amounts in healthy males. However, PSA is often elevated in males with benign prostate hyperplasia (BPH), prostatitis and prostate cancer (161). While PSA measurements can distinguish healthy patients from diseased patients, elevations in PSA do not distinguish between benign and malignant conditions (160). Overall, the utility of PSA screening for prostate cancer is highly controversial, due to limited conceived benefits, the risks of therapeutic intervention upon detection of prostate cancer, and uncertainties regarding the screening strategy (329). These factors have lead to the argument that PSA screening has resulted in overdiagnosis and overtreatment of prostate cancer (330). Recent evidence suggests that factoring in site-specific glycoform composition may improve the PSA screen in distinguishing malignant prostate cancer from benign conditions. PSA glycans from patients with BPH consist predominantly of complex biantennary glycans with various degrees of core fucosylation and sialylation. In contrast, PSA glycoforms from patients with prostate cancer exhibit elevated multiantennary complex

glycoforms with altered sialylation, fucosylation, and LacdiNAc groups (160,161,331). Thus, utilizing absolute PSA levels can identify diseased patients and subsequent analysis of the glycosylation patterns can distinguish malignant from benign conditions.

1.3.4.2.3 Carbohydrate Antigen 19-9 (CA19-9)

CA19-9 (sialyl Lewis A, SLeA), a glycan motif initially identified by Koprowski et al. in a colorectal cancer cell line, is currently the only FDA approved blood test for the management of pancreatic cancer (332,333). It is now recognized that the CA19-9 antigen is present on glycoproteins and glycolipids (334,335). CA19-9 is composed of four monosaccharides, consisting of a Gal and GlcNAc residue with a terminal α2,3-linked sialic acid on the galactose and an α1,4-linked fucose on the GlcNAc. In contrast, SLeX differs in regards to the fucose linkage. The synthesis of both structures are mediated by α2,3 sialyltransferases and α1,3/4 fucosyltransferases, with various degrees of specificity between SLeA and SLeX synthesis (336). CA19-9 is thought to play an important role in disease metastasis through the interaction of the antigen with endothelium E-selectin (337). This claim is evidenced by the abolished interaction of cancer cells with E-selectin following [1] treatment with antibodies against SLeA, [2] the exogenous addition of mucins containing CA19-9 to compete for E-selectin binding, and [3] inhibition of CA19-9 expression (338–340).

While CA19-9 is used as a blood test for pancreatic cancer, the antigen is not used clinically to diagnose pancreatic cancer. Low sensitivity and specificity of CA19-9 in the prediction of pancreatic cancer are exacerbated by the low prevalence of pancreatic cancer, rendering CA19-9 ineffective as an early stage diagnostic biomarker (333). The specificity of CA19-9 is hindered by diseases other than pancreatic cancer, many of which are benign, that express elevated CA19-9 levels; including biliary obstruction, thyroiditis, diabetes, and chronic pancreatitis (341–343). The sensitivity of CA19-9 as a pancreatic cancer biomarker is limited by the fact that between 5-15% of the population do not express fucosyltransferase 3 and are

therefore incapable of expressing CA19-9 (333,343). Efforts to investigate CA19-9 as a screening marker in asymptomatic individuals yield a dismal predicative value positive of .5-.9%, due to a large percent of false positive identifications (344,345). The clinical utility of CA19-9 for the diagnosis of symptomatic patients is also low (346).

While not an ideal diagnostic marker, pre-operative and post-operative serum CA19-9 levels can be useful in assessing cancer stage, resectability, prognosis, response to chemotherapy and postoperative recurrence. A study by Kim et al. demonstrated a rise in mean pre-operative serum CA19-9 levels at each cancer stage, starting at 40.05 U/mL for stage IA patients, and rising to 3239 U/mL for stage IV patients (344). Pre-operative CA19-9 serum levels were found to be lower in patients with localized tumors compared to patients with advanced or metastasized tumors (347). Therefore, it is unsurprising that patients with normal pre-operative CA19-9 levels have a higher 5-year disease specific survival (DSS) than patients with elevated CA19-9 levels, even when comparing patients with similarly staged tumors (348). Furthermore a decline in CA19-9 serum levels following treatment is associated with longer overall survival (346).

1.3.4.2.3 Alpha Fetoprotein (AFP)

AFP is a glycoprotein with a mass of around 70 kDa with one N-glycosylation site, typically occupied by a complex biantennary N-glycan (160). AFP expression is abundant in the human fetus, but levels decrease rapidly during fetal development and after birth (326). Therefore, the detection of AFP in adults is used as a marker for hepatocellular carcinoma (HCC). However, elevated levels of AFP are also observed in benign liver diseases, limiting the use of AFP in distinguishing malignant from benign conditions. An additional limitation of AFP as a biomarker of HCC is the low sensitivity for early stage cancer (349). Efforts to improve the utility of AFP have focused on differential glycosylation of AFP, where α1,6 core fucosylated AFP outperforms AFP alone as a biomarker for HCC (350–352). AFP can be fractionated into three groups using lens culinaris agglutinin (LCA) lectin to provide information of the underlying

cause of AFP elevation; these being AFP-L1, AFP-L2, and AFP-L3. AFP-L3 elevation is linked to malignancy, while AFP-L1 and AFP-L2 are linked to non-malignant liver disease and yolk sac tumors, respectively (160). Other studies have suggested that AFP sialylation can distinguish malignant from benign liver diseases, providing another avenue to enhance biomarker specificity (161). The use of AFP as a biomarker is the best example of how proteomic and glycomic data can be combined to diagnose a cancer.

1.4 Aberrant Glycosylation in Pancreatic Cancer

1.4.1 Biological Factors that Promote Aberrant Glycosylation in Pancreatic Cancer

Pancreatic cancer, like many other cancers, has been characterized as a glycolytic tumor based off a reliance on glycolysis for energy production, as opposed to oxidative phosphorylation, which enables rapid energy production in the hypoxic tumor microenvironment (353,354). Anaerobic glucose metabolism in pancreatic cancer relies on an upregulation of glucose transporters (GLUTs) and glycolytic enzymes including hexokinase II (HK2), phosphofructokinase I and II (PFK1, PFK2), and pyruvate kinase (PKM2) (354). Even surrounding, well oxygenated PDAC cells can utilize lactate produced during glycolysis to promote cell proliferation, as the surrounding cells express elevated levels of the lactate uptake enzyme, monocarboxylate transporter I (MCT1) (353). Pancreatic cancer cells also exhibit a glutamine addiction, in which glutamine can be metabolized to glutamate by glutaminases (GLS1 and GLS2) as a major source of energy. In fact, cells cultured in the glutamine-deficient media exhibit decreased cell viability and proliferation (353). The metabolism of glucose and glutamine both converge upon the hexosamine biosynthetic pathway (HBP) and increase the presence of monosaccharides that can ultimately be incorporated into glycans. The need for glycolytic energy production in pancreatic cancer is perpetuated by the dense stroma and ECM accumulation resulting in a desmoplastic region that surrounds the cancer promotes hypoxia (355). For cancer

in general, more pronounced hypoxic conditions have been linked to metastatic spread, resistance to therapy, and poor prognosis. These trends that hold true in pancreatic cancer, as there is a correlation between hypoxia and reduced survival (355). In addition to promoting glycolysis, hypoxia transcriptionally upregulates glycosyltransferase and sugar transporter expression in cancer cells through the hypoxia-inducible factors (HIFs) (356). Hypoxia has been directly linked to increased glutamine-fructose-6 phosphate aminotransferase I and II (GFPT1 and GFPT2) expression in cell lines (353). Additionally, hypoxia plays a role in cancer invasion and metastasis by promoting the expression of the selectin analogs, SLeA and SLeX, in malignant cells by upregulating fucosyltransferase (FUT7) and sialyltransferase (ST3Gal1) expression (357). Sialin expression, a transporter for NeuGc and NeuAc, is also induced by hypoxia, providing one potential mechanism of the incorporation of non-human NeuGc in malignant cells (211,358). GBPs, such as Galectin-1, are also observed to be elevated in hypoxic tissues.

In addition to altering the glycan biosynthetic machinery, the hypoxic environment in the tumor provides a selective pressure that favors the growth of the most aggressive cells that are capable of surviving in conditions with limited oxygen and nutrients (355). These conditions promote the expansion of cells with mutations in oncogenes and tumor suppressors, such as *KRAS* and *TP53* mutations, respectively (359,360). These mutations are also interconnected with the HBP. Mutated KRAS increases GLUT1, HK1, HK2, PFK1, and lactate dehydrogenase A (LDHA) expression, all of which impact glycosylation by directing glycolytic intermediates to the HBP (40). Extinction of mutated KRAS in cell or xenograft models results in downregulation of a variety of metabolic processes and inhibited tumor grown. The inhibitory effect was directly attributed to diminished glycosylation, as the knockdown of GFPT1 exhibited the same effects on tumor growth (361). In addition to O-glycosylation, activating mutations of RAS oncogenes have also been associated with N-glycan branching.

Given the broad scope of aberrant glycosylation in PDAC, efforts have been directed at identifying the source of aberrant glycosyltransferases that play a role in tumor progression. Specific alterations of glycans structures can, but do not always, coincide with alterations in the expression of glycosyltransferases (210,217). CA19-9 expression is the most well defined marker for pancreatic cancer, so efforts have been directed at identifying the enzymes culpable for the elevated expression of the antigen. Sialyltransferases and fucosyltransferases responsible for sialyl Lewis structure synthesis have been reported to be elevated in pancreatic cancer, including α2,3-sialyltransferases, ST3Gal III, ST3Gal IV, and ST3Gal VI (336). Ectopic expression of ST3Gal III and ST3Gal IV has been linked to an increase in expression of sialyl Lewis antigens and a reduction of α2,6-linked sialic acids, resulting in increased cell motility, reduced adhesion, and increased metastasis in mouse models (336,362,363).

1.4.2 Protein and Glycoprotein Alterations in Pancreatic Cancer

Early efforts to identify proteins that are aberrantly expressed in pancreatic cancer performed comparative studies on whole tissue sections with cancerous and non-diseased regions. Pancreatic cancer cells are often surrounded by stromal cells, which are capable of occupying up to 90% of the total tumor volume, so the overall conclusions did not definitively identify differences between tumor and non-tumor tissue. However, these experiments did provide some mechanisms for cancer progression. There is a disease-related variation in proteins responsible for tumor growth, migration, invasiveness, metastasis and ECM interactions, highlighting the importance of the cancer cell interactions with the stroma and ECM during tumor progression (364,365). Proteins involved in cytoskeletal or ECM regulation, such as annexin A2, cathepsin B, cathepsin D, fascin, gelsolin, and plectin-1, were upregulated in the tumor tissue, providing mechanisms for tumor cells to degrade of the ECM and enter the bloodstream (365–367). Overall, these studies identified a larger number of proteins elevated in the tumor compared to the matched non-tumor tissues, but elevated expression could not be directly linked to the stromal or

epithelial cells (365,366,368). Digestive enzymes, including amylases, peptidases, lipases and elastases, were the main category of proteins that were downregulated in the tumor tissue (366,369). From these initial experiments, proteins of interest were selected for further analysis. Pan et al. evaluated gelsolin concentration as a potential biomarker for pancreatic cancer by SRM analysis, and determined that serum levels of gelsolin to have an AUC = .89 for distinguishing PDAC from pancreatitis and control serum (370). In another study, Bausch et al. found that IHC staining of plectin-1 was able to identify late-stage PanIN lesions and metastatic sites (367). In addition to proteins that regulate cytoskeletal or ECM regulation, proteins that play a role in cell-cell and cell-matrix interaction were also elevated in PDAC and late stage PanIN lesions; including galectin-1, galectin-3, integrin β1, actinin-4, laminin β1, and vimentin (365,366,368,371). The identification of altered proteins in PanIN lesions and full-blown PDAC implicates these proteins in early stages of tumor progression, through their role in cell motility and cell interaction (368,371). Similar studies tabbed versican and lumican as PDAC-elevated proteins that promote metastasis, with plasma levels of lumican achieving an AUC = .94 for differentiating PDAC from control serum by SRM analysis (370). In order to mitigate the role of stromal proteins in the analysis, laser capture microdissection (LCM) was used to isolate pancreatic cancer cells. Downstream proteomics identified S100A6 and annexin III to be elevated in tumor cells, while trypsin and LDH were elevated in non-tumor cells (372). The analysis of serum and biofluids has identified protein families that are consistently observed in pancreatic cancer. Mucin and CEACAM proteins are predominantly expressed in cancer specimen. Additionally, both proteins have been recognized to act as CA19-9 carriers. Research has identified MUC1, MUC2, MUC4, MUC5 and MUC16 as proteins that can be used for the diagnosis or prognosis of diseased pancreas tissue (15,373–375). These proteins have been linked to tumor progression, altered adhesion, and the promotion of perineural invasion. MUC4 stabilizes HER2 expression on the cell surface, which promotes downstream signaling through

FAK and p42/44 (375). MUC1 interacts with Siglec-4a to promote perineural invasion (374). Numerous integrins are upregulated in PDAC including α1, α2, α4, α6, and β1 subunits (363,376).

While initial proteomic experiments were important from a disease marker perspective, the data were also useful in elucidating potential driving mechanisms for pancreatic cancer. These studies largely conclude that among all proteins, glycoproteins are overexpressed (364). Gene ontology annotation of the human pancreas N-glycoproteome demonstrates that a majority of pancreas glycoproteins are involved in cell organization, immune response, adhesion and mobility. The deregulation of these processes may be directly related to the tumorigenic process (377). In addition to alterations in glycoprotein expression, the overall level or structure of glycosylation is altered in pancreatic cancer. Numerous potential disease markers have been identified that differ only in glycosylation and not protein abundance. The ratio of glycosylated RNase1 to total RNAase1 has been tested clinically as a serum biomarker for pancreatic cancer (378). The comparison of non-diseased and diseased ductal fluids revealed kininogen-1 and biglycan to exhibit elevated levels of glycosylation while maintaining a similar protein level (283). Alternatively, fibrillin exhibited elevations in protein expression but a reduction in glycan occupancy. The same study found that occupancy rates of different glycopeptides from the same protein varied significantly and independently, allowing unique sites to be concurrently elevated and diminished in diseased samples compared to non-diseased samples. For example, apolipoprotein B-100 had two glycosylation sites that exhibited elevated occupancy and three sites that had reduced occupancy in diseased samples (283).

Unfortunately, there is a large degree of similarity between the pancreatic cancer and pancreatitis proteome, especially when comparing PDAC to severe chronic pancreatitis, which exhibited more dysregulation of the proteome than mild pancreatitis. Among highlighted proteins identified in the earlier global proteome analysis, cathepsin D, integrin β1, plasminogen, versican

and lumican are all over-expressed in both PDAC and pancreatitis compared to control samples (379,380). Factoring information about the protein expression level and associated glycoforms can distinguish malignant from benign conditions. At the protein level, the trend of elevated glycoproteins in PDAC was also observed in pancreatitis. However, efforts to identify glycoproteins that can distinguish PDAC from both benign and non-diseased samples following glycoprotein enrichment have been fruitful. Ingenuity pathway analysis (IPA) of elevated glycoproteins unique to PDAC in patient ductal fluids emphasized the role of the proteins in endothelial and tumor cell movement and interaction. Mucin and CEACAM proteins were represented among the proteins that were only elevated in tumor (283). A similar study identified members of the REG protein family, phospholipase A2 and elastase 2B to be present in only ductal fluid from patients with cancer or premalignant lesions (381). Other proteins were elevated in tumor compared to pancreatitis specimen, including versican, biglycan, haptoglobin, annexin A2, IGFBP-2 and apolipoprotein E (283,364,382). In addition to assessing glycoprotein expression levels, studies have also assessed glycoform structure and occupancy on those proteins. Nie et al. used AAL lectin to enrich fucosylated glycoproteins in an effort to identify individual serum glycoprotein discriminators to distinguish serum from patients with PDAC, no disease, diabetes, pancreatitis, obstructive jaundice, and benign cysts (343). Using LC-MS/MS for initial glycoprotein identification followed by ELISAs and lectin ELISAs for validation, numerous glycoproteins were determined to perform better than CA19-9 in disease discrimination. Importantly haptoglobin, lumican, α-1-antichymotrypsin and α-1-antitrypsin were able to distinguish pancreatic cancer from benign diseases more effectively than CA19-9 (343). Finally, a biomarker panel of proteins consisting of α-1-antichymotrypsin, thrombospondin-1, and haptoglobin showed improved utility and an increased AUC for distinguishing PDAC from cancer (AUC = .95), diabetes (AUC = .89), cyst (AUC = .82), and pancreatitis (AUC = .90) (343). As the expression of these proteins was not highly correlative with CA19-9, adding CA19-

9 into the biomarker panel improved the AUC values. Other assays also showed improved biomarker performance by combining protein and CA19-9 expression (383). Lectin enrichment of α2,6-linked sialic acids from serum samples revealed elevated levels of disialylated biantennary glycans (Hex5HexNAc4NeuAc2) on α-1-antitrypsin in PDAC, but decreased levels of the glycoform in pancreatitis (318). Overall, both conditions had elevated levels of α-1-antitrypsin, so only the glycoform difference can be used to distinguish benign from malignant conditions. The same glycoform was found to be elevated on transferrin in pancreatitis over PDAC (318). Protein expression and glycoform occupancy do not always correlate. While haptoglobin was elevated in pancreatic cancer serum, glycoform differences were only observed in pancreatitis (318,382). Given the divergence between the proteome and the glycome, the combination can be used to enhance the utility of biomarkers. Li et al. evaluated a multiplexed bead assay directed at quantifying specific proteins and the corresponding levels of α2,6-linked sialic acid in order to distinguish PDAC from pancreatitis and control serum. This study concluded that serum levels of amyloid P were significantly lower in PDAC than non-diseased serum, while the glycoforms on α-1-β glycoprotein vary significantly between PDAC and pancreatitis (384). The functional role of dysregulated protein glycosylation has also been investigated. Integrins have also been implicated in PDAC progression through their modulatory effects on adhesion and invasion. Integrins frequently exhibit higher levels of sialylation, which alters their binding to the ECM and regulates tumor cell migration (336,363).

1.4.3 Pancreatic Cancer Glycomics

Most experiments assessing glycosylation changes in pancreatic cancer have utilized serum samples. While noninvasive to obtain, serum glycoproteins often originate in areas other than the malignant site, reducing the probability of cancer identification based off of global glycosylation changes (327). Broad changes in PDAC serum glycosylation have been assessed by interaction with specific lectins and have identified a notable increase in terminal α-fucosylation

in serum from PDAC patients compared to non-diseased and benign controls (343). Additional lectin array data has identified a variety of other glycan patterns that are elevated in PDAC, including the CA19-9 antigen, complex/branched glycans, O-GlcNAc modifications, β-GalNAc residues, and α2,6-linked sialic acids (373). However, limited studies have assessed individual glycoform variations in pancreatic cancer. A broad glycomic analysis of N-glycans from serum and ductal fluids found that only serum glycans could distinguish diseased from non-diseased patients, and that the glycans could also provide information on metastatic progression (381,385). Serum from PDAC patients exhibited an overall higher amount of glycosylation, specifically in regards to elevated branching, fucosylation, and sialylation. On the other hand, the level of high-mannose and hybrid glycans did not differ in PDAC serum compared to control serum (385). Elevated levels of LacdiNAc groups have also been observed in pancreatic cancer (331). Glycomic analysis of ductal fluid was marred by glycoform heterogeneity but did identify driver glycans that could cluster glycans into specific groups, which could then be used for diagnostic classification. However, this study was done on an extremely limited sample set, so it remains to be seen how this clustering method performs over larger sample sizes (381).

1.5 Broad Overview

To summarize, the dismal outcomes of pancreatic cancer necessitate in-depth research to identify early stage disease markers and evaluate mechanisms involved in disease progression. Aberrant glycosylation, a hallmark of cancer, has been linked to pancreatic cancer through the regulation of glycosylation synthetic machinery and elevated expression of CA19-9. Apart from potential roles of CA19-9 and sialylated glycoconjugates in adhesion and metastasis, the role that aberrant glycosylation plays in the tumorigenic process is undefined. Cancer-related trends in glycosylation have been observed, but knowledge of individual glycoforms that are encompassed within these trends is absent. Efforts to study pancreatic cancer glycosylation are further hindered by the low prevalence of pancreatic cancer, which limits the number of samples available to study

While alterations of glycoforms can be used to identify the presence of a disease, it is perhaps of more use to identify aberrantly expressed glycoproteins that are indicative of pancreatic cancer, both in terms of protein abundance and glycoform occupancy on the protein. Linking aberrantly expressed glycoforms, including the CA19-9 antigen, to the underlying protein in pancreatic cancer offers the potential to identify pancreatic cancer biomarkers. Both glycomic and glycoproteomic experiments initiated in tissue specimen can be further assessed in biological fluids as non-invasive disease markers. While glycomic and glycoproteomic research related to pancreatic cancer is lacking, the advent of novel methodologies to improve throughput, reduce sample handling, and overcome limitations in detection, can advance our understanding of the pancreatic cancer glycoproteome and identify putative disease markers for further validation. The relevance of technologies such as MALDI-IMS, glycan derivatization, click chemistry, and intact glycopeptide analysis will be outlined in this dissertation in relation to pancreatic cancer glycomics and glycoproteomics.

Chapter 2: Hypothesis

The overarching goal of this research project is directed at developing and applying novel mass spectrometry-based technologies for the identification of glycan and glycoprotein disease markers for pancreatic cancer. The methods developed and adapted towards this goal include: [1] the development of a MALDI-IMS approach to identify N-glycoforms with retention of their spatial distribution across tissue sections, [2] adaptation of a novel N-glycan derivatization strategy that stabilizes sialic acid residues and differentiates their linkages based on an induced mass difference, [3] azido-sugar metabolic labeling of sialic acids followed by protein enrichment and identification by LC-MS/MS, and [4] intact analysis of glycopeptides using a product-dependent fragmentation strategy. The importance of applying these novel technologies for pancreatic cancer is based on the fact that pancreatic cancer consistently ranks among the most deadly cancers, due to delayed diagnosis, rapid disease progression, and limited therapeutic efficacy. The only curative treatment for pancreatic cancer is disease resection, and this option is available to only a small subset of patients whose disease is not locally invasive or metastasized at the time of diagnosis. Identification of early stage disease markers for pancreatic cancer offers the potential to improve clinical outcome by increasing the number of patients who can undergo disease resection.

Glycoproteins and the glycoforms present on those proteins represent a promising avenue for early detection biomarker discovery research, given the established evidence linking aberrant glycosylation to pancreatic cancer. In addition to elevated levels of CA19-9 in pancreatic cancer, evidence has also shown that pancreatic cancer is associated with alterations in glycosylation biosynthetic machinery, broad structural changes in N-linked glycans, and robust changes in glycoprotein expression levels. However, the low prevalence of disease and even lower number of clinical specimen available, have hindered the ability to identify disease markers. Furthermore, techniques directed at studying glycobiology are both labor and time intensive, and require large amounts of sample material. While glycobiology research has lagged behind proteomics and

genomics in recent years, new analytical technologies offer the potential to produce profound improvements in the sector. Therefore, I **hypothesize** that the application of analytical techniques that characterize the glycome of the tumor and adjacent regions, and a subsequent analysis of the dysregulated glycoproteome, will identify molecular events associated with the development and progression of pancreatic cancer. A premise of this hypothesis is that these alterations in pancreatic cancer glycosylation are conserved across multiple patients. This hypothesis is tested by the following specific aims:

Aim 1: Develop a method to simultaneously indentify and localize N-glycans on tissue sections, while maintaining information related to relative abundance

Approach

A novel method to spatially profile N-linked glycans on biologically relevant tissue sections will be developed for MALDI-IMS to enable the identification, abundance and localization of N-linked glycans in tissue sections. While most MALDI-IMS approaches utilize fresh/frozen tissue sections, the method will be optimized to include formalin-fixed paraffin-embedded (FFPE) tissue blocks. Initial characterization will rely on mass accuracy of released glycans, but secondary characterization strategies, including glycan fragmentation and derivatization, will confirm glycan compositions. Once optimized, this method will be used to identify diagnostic and prognostic glycan alterations directly in the clinical tissue blocks. Towards this end, the use of tissue microarrays (TMAs) will improve assay throughput and reduce experimental acquisition times.

Rationale

The intent of Aim 1 is to improve glycobiology methodologies relating to the study of N-glycan localization. In 2012, the U.S. National Research Council laid out a series of goals to advance the field of glycobiology research; one being to develop imaging methods to assess glycan localization (386). The current method for localizing N-glycans in tissue sections relies on staining with broad affinity lectins or antibodies specific to one glycan antigen. Therefore, a global profiling of N-glycans using traditional approaches requires serial staining of multiple slides, requiring the use of excessive materials (i.e. tissue slices, lectins, and antibodies), and still does not definitively identify the glycan composition. MALDI-IMS is a developing technique that has been used to study the biological localization of drugs, lipids, metabolites and proteins in tissue slices. MALDI-IMS following enzymatic digestion has already demonstrated promise in the field of bottom-up protein imaging. By adapting a similar workflow, we aim improve throughput and dimensionality of N-glycan imaging by globally profiling all N-glycans in a single experiment. PNGaseF will be applied to tissue sections prior to imaging to release glycans at the N-glycan consensus sequence. Following analysis, each slide will be stained with hematoxylin and eosin and processed by a pathologist to enable a comparison of glycan localization with histology. Tissue embedding in FFPE blocks is considered the gold standard for sample preservation, so adapting the imaging method to enable FFPE tissue analysis will expand the scope of studies (387). An additional benefit of optimizing the method for FFPE tissues is the ability to use FFPE cores housed in a TMA blocks to look at a large number of samples in a single imaging run.

Aim 2: Use higher-throughput MALDI imaging approaches to identify N-glycan biomarker panels for the identification of pancreatic cancer

Approach

The developed MALDI-IMS platform will be applied to the study of pancreatic cancer by profiling N-glycans from pancreas tissues and TMA blocks containing patient matched tumor and non-tumor tissue cores. Initial disease markers for pancreatic cancer will be discovered by higher-throughput N-glycan profiling of 6 TMAs with tissue cores from 76 patients. Data will be processed by supervised machine learning algorithms to identify panels of N-linked glycans that are capable of distinguishing tumor from non-tumor cores. The utility of individual glycan and glycan panels can be validated in high resolution imaging experiments of entire pancreatic cancer tissue blocks. Given the heterogeneous nature of tumors and the presence of a significant percentage of the population that is unable to express CA19-9 antigen, we expect biomarker panels to improve the overall disease marker utility.

Rationale

The goal of our work in Aim 2 is to assess the ability of N-glycans to serve as disease markers for pancreatic cancer, a facet ignored by most glycosylation-centered pancreatic cancer studies. The observation of specific glycoproteins has been used to detect early stage pancreatic lesions, suggesting that aberrant glycans may have a similar role. Additionally, the development of a glycan disease-marker panel can improve the sensitivity and specificity of disease markers compared to CA19-9. To identify individual glycans and glycan panels that can be used as biomarkers for pancreatic cancer, we aim to utilize the method developed in Aim 1 to perform discovery phase experiments in TMAs and tissue sections. Tissues are ideal for discovery phase disease marker studies due to reduced sample complexity compared to serum, as molecules are obtained directly from the organ of interest and are present in higher levels. An added benefit of MALDI-IMS is the ability of the technique to link glycan structures with histology. Pancreatic

cancer is a very complex disease that requires input from cancer cells and surrounding stromal cells, an interaction that would be missed by traditional biochemical approaches (95).

Aim 3: Link aberrant glycosylation to modified proteins and develop an integrative glycoproteomics method that incorporates glycoform and glycoprotein expression in a single experiment

Approach

Additional disease markers will be identified by leveraging information relevant to individual protein expression and glycoform occupancy/heterogeneity. Due to the inherent complications of glycoprotein analysis, novel methods will be adapted to the study of the pancreatic cancer glycoproteome. In one approach, putative carriers of the CA19-9 antigen will be assessed using a targeted glycoprotein enrichment procedure specific for sialic acids. In this approach, cells with known CA19-9 expression levels will be treated with an azide-labeled sialic acid analog. Alterations at the protein level may identify common trends associated with CA19-9 expression, and selected proteins can be further evaluated as CA19-9 carriers. Additionally, this approach will catalogue sialylated glycoproteins in pancreatic cancer. In another approach, an intact glycopeptide enrichment step followed by a product-dependent fragmentation of selected glycopeptides will be used to simultaneously characterize both glycan compositions and peptide sequences. This experimental workflow will be developed with protein standards and then attempted on complex biological mixtures (i.e. tissue, serum, cell lysates). Glycoconjugate enrichment will be performed at the peptide level though unbiased enrichment strategies that rely on the physiochemical properties of the glycan.

Rationale

The ultimate goal of Aim 3 is to identify aberrations in both glycans and their corresponding glycoproteins that can be used as disease markers for pancreatic cancer in a clinical setting. While it is doubtful that MALDI-IMS will be used clinically for biomarker classification, derivative experiments based off of identified glycan or glycoproteins can be assessed by immunohistochemistry, ELISA, lectin-ELISA, or MRM analysis. Clinically, secreted glycoproteins are successful non-invasive biomarkers for a variety of diseases, including hepatocellular carcinoma, prostate cancer, colon cancer and breast cancer (271). However, these glycoproteins are often elevated in benign diseases, resulting in false diagnosis and additional screens. Recent evidence has suggested that monitoring site-specific protein glycosylation can improve the biomarker, an example being elevated core fucosylated glycoforms of AFP in hepatocellular carcinoma compared to benign liver diseases (352). In pancreatic cancer, certain glycoproteins have displayed improved sensitivity and specificity over CA19-9 in distinguishing between pancreatic cancer, normal pancreas tissue, and other diseases (i.e. pancreatitis or pancreatic cysts). However, these trends are observed in small subsets of data, but require further validation in larger trials (343). Targeted approaches, focusing on glycosylation motifs will be used to assess specific aberrations at the glycoprotein level. However, the click chemistry approach utilized in this aim does not retain glycans on the protein. Mapping specific changes in glycosylation unique to pancreatic cancer may require glycoproteins to be isolated in order to reduce sample complexity, prior to glycoproteomic analysis. The workflow for intact glycopeptide analysis will be optimized using glycoprotein standards.

Chapter 3: N-Glycan Imaging by MALDI-IMS

3.1 Introduction

MALDI-IMS can simultaneously map the molecular distribution and relative abundance of hundreds of analytes on a single tissue section. Furthermore, the distribution can be directly linked to cellular histology on serial tissue sections. While MALDI-IMS has not been applied to the study of glycosylation, the technique has been used to profile the distribution of lipids, metabolites, drugs, proteins and peptides within tissue sections. Given that MALDI-MS is a powerful technique to study N-glycosylation, albeit usually following glycan derivatization and enrichment, the application of PNGaseF to tissue sections should release glycans and enable analysis by MALDI-IMS. Glycosylation is a dynamic PTM that is dysregulated in response to malignant transformation and promotes cancer progression and metastasis. Mapping glycosylation in correlation with disease histology could identify cancer related glycans and implicate glycosylation-related processes in disease progression. In comparison to other classes of molecules, defining the underlying monosaccharide composition of N-linked glycans is relatively straightforward, as N-glycans possess a common core structure and elongation proceeds through a limited number of well characterized processes. Current approaches to define glycan localization are limited to broad affinity lectins and a small number of antibodies directed against glycan antigens, both of which enable the localization of structural classes of glycans but not identification of individual structures. Furthermore, glycans bound to proteins should be accessible in FFPE tissue blocks. As FFPE fixation represents the gold standard of tissue preservation, adaptation of an imaging approach to profile N-glycans on these sections is extremely relevant. From a biomarker discovery standpoint, individual tissue cores from different patients can be profiled simultaneously in a higher-throughput TMA based format. A targeted MALDI-IMS approach to localize N-linked glycans in relation to histopathology can serve as a platform for identifying novel disease related targets for biomarker and therapeutic applications.

3.2 MALDI-IMS Workflow For N-Glycan Imaging

Manuscript published in *Analytical Chemistry*, October 2013 85(20): 9799-9806. TWP performed all of the mass spectrometry experiments and writing of the manuscript. RRD contributed intellectually to the manuscript. ASM performed the HPLC analysis of extracted glycans that were provided by TWP.

A MALDI Imaging Mass Spectrometry Workflow for Spatial Profiling Analysis of N-linked Glycan Expression in Tissues

Thomas W. Powers[1], **E. Ellen Jones**[1], **Lucy R. Betesh**[2], **Patrick Romano**[2], **Peng Gao**[1], **John A. Copland**[3], **Anand S. Mehta**[2], **and Richard R. Drake**[1]

[1]Department of Cell and Molecular Pharmacology and Experimental Therapeutics and MUSC Proteomics Center, Medical University of South Carolina, 173 Ashley Avenue, Charleston, SC, 29425 USA

[2]Department of Cancer Biology, Mayo Clinic Comprehensive Cancer Center, 4500 San Pablo Road, Jacksonville, FL 32224, USA

[3]Drexel University College of Medicine, Department of Microbiology and Immunology and Drexel Institute for Biotechnology and Virology, 3805 Old Easton Road, Doylestown, PA 18902, USA

Corresponding Author:

Richard R. Drake, Ph.D.

Email: draker@musc.edu; FAX (843) 792-0481

3.2.1 Abstract

A new Matrix Assisted Laser Desorption Ionization Imaging Mass Spectrometry (MALDI-IMS) method to spatially profile the location and distribution of multiple N-linked glycan species in tissues is described. Application of an endoglycosidase, peptide N-glycosidase F (PNGaseF), directly on tissues followed by incubation releases N-linked glycan species amenable to detection by MALDI-IMS. The method has been designed to simultaneously profile the multiple glycan species released from intracellular organelle and cell surface glycoproteins, while maintaining histopathology compatible preparation workflows. A recombinant PNGaseF enzyme was sprayed uniformly across mouse brain tissue slides, incubated for two hours, then sprayed with 2,5- dihydroxybenzoic acid matrix for MALDI-IMS analysis. Using this basic approach, global snapshots of major cellular N-linked glycoforms were detected, including their tissue localization and distribution, structure and relative abundance. Off-tissue extraction and modification of glycans from similarly processed tissues and further mass spectrometry or HPLC analysis was done to assign structural designations. MALDI-IMS has primarily been utilized to spatially profile proteins, lipids, drug and small molecule metabolites in tissues, but it has not been previously applied to N-linked glycan analysis. The translatable MALDI-IMS glycan profiling workflow described herein can readily be applied to any tissue type of interest. From a clinical diagnostics perspective, the ability to differentially profile N-glycans and correlate their molecular expression to histopathological changes can offer new approaches to identifying novel disease related targets for biomarker and therapeutic applications.

3.2.2 Introduction

The majority of proteomic and metabolomic analytical techniques require the macro- or microdissection and subsequent extraction of analytes from the target tissue. This process leads to loss of the spatial distribution and associated histopathology of the tissue. A novel but maturing

technology, MALDI imaging mass spectrometry (MALDI-IMS), has been used to generate two- and three-dimensional molecular maps of hundreds of analytes directly from the surface of a tissue section, allowing the display of the relative abundance and spatial distribution of individual analytes (388–392). The distribution of the analytes are also readily linkable to molecular histology and pathology data from the same tissues (393,394). To date, most applications of MALDI-IMS have focused on profiling of proteins, lipids and drug metabolites in tissues, but the technique has not yet been defined for analysis of complex carbohydrates, typified by N-linked glycans (138,143,395–399). Glycosylation is a major post-translational modification to proteins critical in regulating protein folding and vesicular transport, cell-cell communication and adhesion, immune recognition and other extracellular functions (150,400,401). N-linked glycans are specifically attached to asparagine residues in proteins via a conserved amino acid motif of N-X-S/T, where X represents any amino acid except proline, and represent a diverse but biosynthetically definable group of carbohydrate structures ranging generally from m/z = 1000 - 5000 in size (150). Because MALDI-TOF analysis is one of the most robust and well established methods for profiling multiple species of N-linked glycans, imaging of glycans directly on tissue by MALDI-IMS should be feasible (402,403).

Using a recombinant source of peptide N-glycosidase F, which allowed an abundant supply of enzyme for optimization, and adaptation of a molecular spraying method developed for on-tissue protease digestions, a method workflow for MALDI-IMS of released N- glycans has been developed that maintains their spatial distribution in frozen tissue specimens (395,404). Combinations of permethylation derivatization, 2-aminobenzoate-modified normal phase HPLC separations, glycan standards, and existing structural database resources were used to confirm glycan release and initial structural determinations (238,405,406). Examples of the method development and verification workflows for robust on-tissue N-linked glycan profiling by MALDI-IMS are presented for mouse brain and human kidney tissues.

3.2.3 Materials and Methods

3.2.3.1 Materials

The glycan standard A2 and sialidase S were obtained from ProZyme (Hayward, CA). Asialofetuin glycoprotein, 2,5-dihydroxybenzioc acid (DHB), trifluoroacetic acid, sodium hydroxide, dimethyl sulfoxide (DMSO), and iodomethane were obtained from Sigma-Aldrich (St. Louis, MO). HPLC grade methanol, ethanol and water were obtained from Fisher Scientific. ITO slides were purchased from Bruker Daltonics (Billerica, MA) for MALDI-IMS experiments.

3.2.3.2 Tissue Specimen

Mouse brains were excised from four euthanized C57BL/6 mice and snap frozen. Mice were housed in an Institutional Animal Care and Use Committee-approved small animal facility at MUSC, and brains were harvested as part of approved projects. Whole excised brains were placed in plastic weigh boats and rapidly frozen at the vapor phase interface of a liquid nitrogen containing dewar. A human kidney tissue was obtained from an Institutional Review Board approved study from the Mayo Clinic (Jacksonville, FL) evaluating the molecular changes associated with matched patient non-tumor and clear cell renal cell carcinoma tissues. A slice of kidney from a non-tumor region was selected for use in this current study. No personal identifier information was provided to the laboratory investigators. This particular tissue section represents a distal section of kidney where there was no evidence of tumor. Coronal sections of mouse brains (10 microns) and the human kidney (10 microns) were prepared using a Thermo Microm HM550 cryostat and stored at −80°C. For each section analyzed, a serial section was collected for histological analysis and staining with hematoxylin and eosin (H&E).

3.2.3.3 Recombinant Peptide N-Glycosidase F Expression

The entire Peptide N-Glycosidase F (PNGase F) gene (P21163.2) from the genome of *Flavobacterium meningosepticum* was cloned, expressed and purified into the T7 expression

vectors pET 29-b (Novagen) and pQE-T7 (Qiagen). This insertion results in the addition of the sequence for an in frame histidine tag at the C-terminal end of the PNGase F gene sequence. The recombinant PNGase F HIS-tagged construct (rPNGase F) was transformed into bacterial strain BL21 Star (DE3) that carries the gene for the T7 RNA polymerase under control of the *lac*UV5 promoter and allows for high level isopropyl-beta-D-thiogalactopyranoside (IPTG) inducible expression of gene products from T7 expression vectors such as pET and pQE. Bacterial transformation and cell culture growth conditions were done following manufacturers protocol (Invitrogen). Cells were harvested by centrifugation and cell pellets were washed with 1X phosphate buffered saline (PBS) containing protease inhibitors (Sigma*Fast* EDTA-free). Total cellular protein lysates were made using an Avestin C5 high pressure homogenizer. FPLC purification methods for the rPNGase F histidine tagged protein used Ni-NTA (Qiagen) and IMAC HisTrap HP (GE Healthcare) columns. Bacterial cell lysate from IPTG induced cultures was loaded onto to the column in binding buffer; 50 mM $NaHPO_4$, 1M NaCl (pH 7.0) and bound histidine tagged rPNGasE F was washed and eluted using an imidazole step gradient (250mM to 800mM imidazole) in binding buffer. Purified enzyme was dialyzed and stored in PBS buffer at a concentration of 1mg/mL, analogous to concentrations found in commercial sources of this enzyme.

3.2.3.4 MALDI Imaging Mass Spectrometry

Sectioned mouse brain tissue samples were mounted on ITO coated slides, and desiccated at room temperature for 20 minutes. A sequential ethanol wash was done in 70% ethanol for two minutes, two times, and one wash in 100% ethanol for two minutes, then a final desiccation at room temperature for 10 minutes. An ImagePrep spray station (Bruker Daltonics) was used to coat the slide with a 0.2mL solution containing 20µL of 1mg/mL PNGaseF stock in water. This solution was sprayed by the ImagePrep using settings originally optimized for spraying of trypsin that result in minimal volumes and retention of spatial distribution (395,403). If sialidase was

used, 0.1 units of enzyme were diluted in 200µL water and sprayed following PNGaseF application using the same ImagePrep settings. Control tissue slices were blocked with a glass slide during the spraying process. Following application of glycosidase, slides were then incubated at 37°C for 2hrs in a humidified chamber, then dried in a desiccator prior to matrix application. 2,5-Dihydroxybenzoic acid (DHB) matrix at a concentration of 0.2M in 50% MeOH 1% TFA was sprayed on to the slide using the ImagePrep for positive ion mode analysis. Spectra were acquired across the entire tissue section on a SolariX 70 dual source 7T FTICR mass spectrometer (Bruker Daltonics) to detect the N-glycans (m/z = 690–5000) with a SmartBeam II laser operating at 1000 Hz, a laser spot size of 25µm and a raster width of 125µm unless otherwise indicated. For each laser spot, 1000 spectra were averaged. Images of differentially expressed glycans were generated to view the expression pattern of each analyte of interest using FlexImaging 3.0 software (Bruker Daltonics). Following MS analysis, data were loaded into FlexImaging Software focusing on the range m/z = 1200–4500 and reduced to 0.98 ICR Reduction Noise Threshold. Glycans were identified by selecting peaks that appeared in an average mass spectrum from the tissue that received PNGaseF application but did not appear in the control tissue. Observed glycans were searched against the glycan database provided by the Consortium for Functional Glycomics and known glycans that exist in mouse brain (407).

3.2.3.5 Permethylation and Normal Phase HPLC of Tissue Extracted N-Glycans

PNGaseF sprayed mouse brain tissue slides were incubated for 2hr at 37°C; 50µL water was applied on top of the tissue and incubated for 20 minutes to extract the released native N-glycans. The water was removed from the tissue, then concentrated under vacuum by centrifugation. Permethylation was performed as described, and glycans analyzed by MALDI (238). Masses detected in the permethylation experiments were searched against the permethylated glycan database provided by the Consortium for Functional Glycomics (407). Alternatively, concentrated glycans were labeled with 2-aminobenzamide (2-AB) for subsequent

normal phase HPLC analysis (405,406). Glycan identification was made through sequential exoglycosidase digestions as previously described (406).

PNGaseF (0.02mg) sprayed mouse brain tissue slides were incubated for 2hrs at 37°C; 50uL water was applied on top of the tissue and incubated for 20 minutes to remove the released glycans. The water was removed and concentrated under vacuum by centrifugation. Permethylation was performed by dissolving the dried glycans in 0.2μL H_2O, 22μL iodomethane and 50μL dimethylsulfoxide and applied over a sodium hydroxide microspin column as described (238). The permethylated glycans were dissolved in 5μL 50% MeOH/water and spotted with DHB (10mg in 1mL 50% MeOH). N-glycans were detected by averaging 10 MALDI scans at 25 μm laser width, averaging 1000 shots in each acquired scan between m/z = 690-5000. Masses detected in the permethylation experiments were loaded into DataAnalysis 4.0 and searched against the permethylated glycan database provided by the Consortium for Functional Glycomics (407).

Tissue extracted PNGaseF released N-glycans (as above) were prepared for rapid glycan sequencing using previously optimized procedures (405,406). Briefly, the concentrated N-glycans were labeled with 2-aminobenzamide (2-AB) for subsequent normal phase HPLC analysis (406). Glycan identification was made through sequential exoglycosidase digestion as previously described by Guile et al., as well as bovine kidney fucosidase (406). The resulting peaks, separated by time of appearance, correspond to specific glycan structures on the basis of glucose unit values (data not shown). All HPLC analyses were performed using a Waters Alliance HPLC System and quantified using the Millennium Chromatography Manager (Waters Corporation, Milford, MA). Glycan structures were identified by the calculation of the glucose unit value, as previously described, as well as through the comparison to known standards and sequential exoglycosidase digestion[22,23] (405,406).

3.2.3.6 Analysis of Asialofetuin and Glycan Standards in Mouse Brain Tissues

A frozen mouse brain was cored with a disposable 1.5mm tissue coring tool. A quartile of a different brain from a littermate, which contained the same region of tissue cored in the other sample, was homogenized in a glass Dounce homogenizer in 0.5mL water. The goal was to maintain a highly viscous preparation, which was then supplemented with 0.02mL of water containing 1mg purified asialofetuin glycoprotein standard. This solution was added back to the cored brain, and refrozen at -80°C. This brain was subsequently prepped for MALDI imaging as described for other brain samples.

3.2.4 Results

A workflow for on-tissue analysis of PNGaseF released N-glycans was developed, as summarized in Figure 7. Initial experiments were done to verify that PNGaseF is active on mouse brain tissues following application by an ImagePrep sprayer, and that the products of the reaction are N-glycans. Two frozen mouse brain tissue slices (10 micron) were mounted on ITO-coated glass slides and dehydrated in sequential ethanol washes. One brain slice was covered by a glass slide to serve as an experimental control, and the other section was sprayed with 20μg PNGaseF. The tissues were incubated for 2hr at 37°C in a humidified chamber. After drying, DHB matrix was applied to the entire slide and MALDI-IMS imaging done for each brain slice. The spectra from the PNGaseF sprayed brain slice revealed multiple novel ions peaks that are consistent with release of N-glycans (Figure 8a). Additionally, these ions displayed regional localization following PNGaseF spray (Figure 8b, 2d, 2f, 2h) but were not detected in the absence of PNGaseF spray (Figure 8c, 2e, 2g). These obvious patterns of differential glycan expression distributions associated with brain substructure is a novel finding, and a representative overlay image of PNGaseF released glycans is shown in Figure 8h.

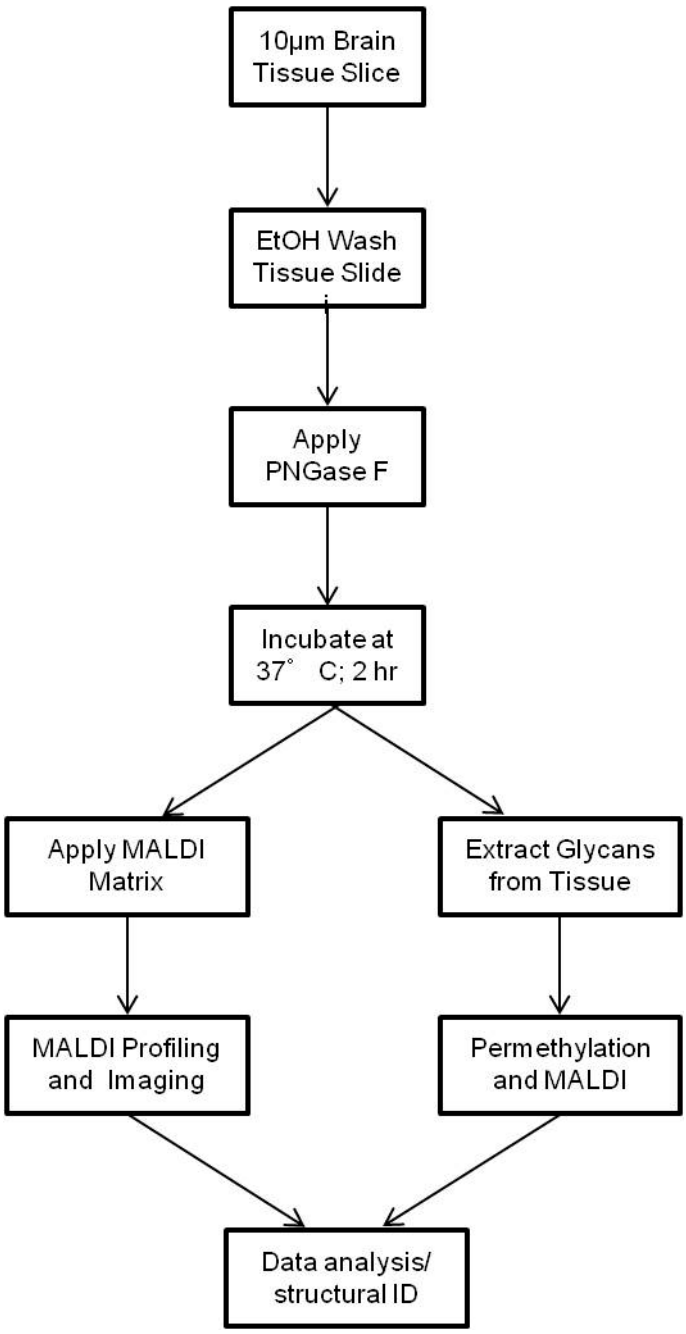


Figure 7. MALDI-IMS of N-Linked Glycans Workflow. A summary schematic of the on-tissue PNGaseF digestion for MALDI-IMS

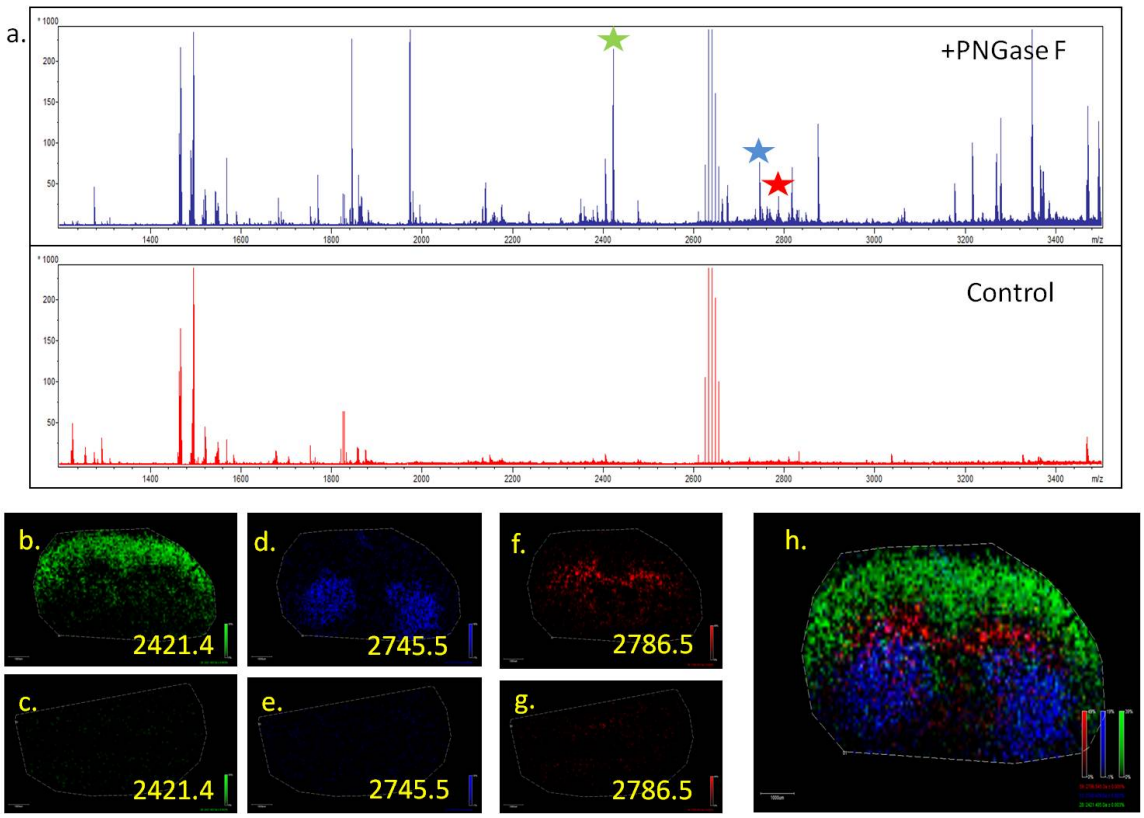


Figure 8. Application of PNGaseF Results in Enzyme-Mediated Generation of Ions. The PNGase F workflow was applied to one mouse brain while the other mouse brain was prepared with an identical workflow without PNGase F. (a). The upper spectrum is from the brain slice incubated with PNGase F and the lower spectrum is from the non-PNGase F control brain slice. Example masses that were detected following PNGase F treatment are shown for three ion peaks that also displayed regional expression distributions in the mouse brain. The m/z = 2421.5 (b and c), m/z = 2745.5 (d and e) and m/z = 2786.5 (f and g) were only detected after PNGase F digestion (panels b,d,f,h). An overlay of these three peaks in the PNGase F digested brain slice is shown in panel h. A raster distance of 75 microns was used per MALDI laser acquisition across the tissues.

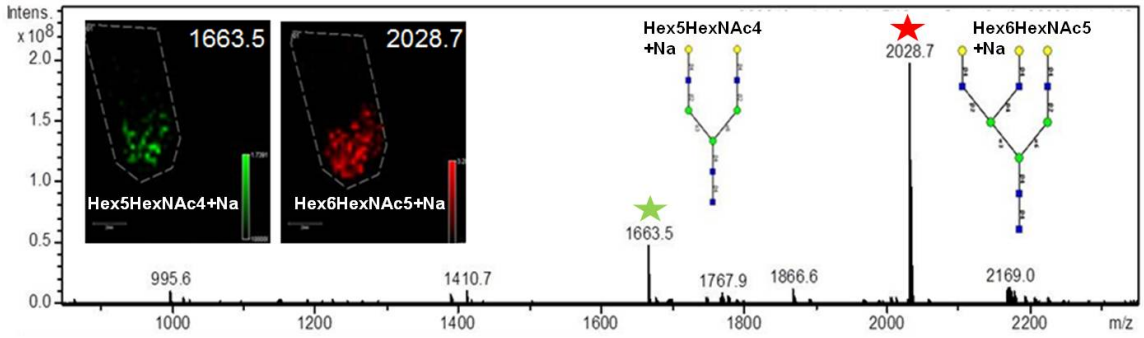


Figure 9. Addition of a Glycoprotein Standard Following PNGaseF Digestion and MALDI-IMS. Asialofetuin was incubated with PNGase F and the native glycans were analyzed by FT-MS. Hex5HexNAc4 + Na (m/z = 1663.5) and Hex6HexNAc5 + Na (m/z = 2028.7) were detected. Asialofetuin was incorporated into a mouse brain and the PNGase F imaging method was performed. The two inset images show the same N-glycans at m/z = 1663.5 and m/z = 2028.7 m/z detected on-tissue following PNGase F digestion.

Purified glycoprotein and glycan standards were added to mouse brain tissues prior to PNGaseF digestion and matrix application, to further verify the on-tissue PNGaseF activity and the ability to detect native N-glycans on tissue. A common purified glycoprotein standard, asialofetuin, was mixed with brain tissue homogenate and added back to the brain for refreezing and MALDI-MS imaging analysis of the tissue. Focusing on two known N- linked glycans of asialofetuin, Hex5HexNAc4 + Na (m/z = 1663.5) and Hex6HexNAc5 + Na (m/z = 2028.7), these glycans were detected following in-solution digest of asialofetuin alone (Figure 9). Following on-tissue PNGaseF digestion, these same glycans were readily detected in the supplemented brain tissue following PNGaseF digestion (inset of Figure 9), but not in the tissue without the enzyme sprayed (data not shown). Furthermore, addition of a glycan A2 standard (m/z = 2223.8, Hex5HexNAc4NeuAc2) directly on mouse brain tissues revealed multiple glycan ion peaks that differed in loss of sialic acids (−1 NeuAc, m/ z = 1954.7; −2 NeuAc, m/z = 1663.5) (Figure 10), a common occurrence in MALDI analysis of sialylated glycan structures, and multiple sodium ion adducts (247). The corresponding MALDI imaging distribution of the standard glycan peaks are shown for each ion in Figure 10. Analysis of these standards further verified detection of N-linked glycans by MALDI-IMS.

Two off-tissue approaches were used to further verify the release of N-glycans and begin to identify specific glycan structures associated with on-tissue native glycan peaks. Brain tissues were digested with PNGaseF on-slide as already described. Released glycans were extracted from the on-slide digestion in water. On-slide extracted glycans were treated with 2-aminobenzoate (2AB) for subsequent normal-phase HPLC analysis, or permethylated for further MALDI analysis and comparison to structure and reference spectra databases (405–407). Reference spectra for these extracted glycans following permethylation and MALDI profiling and normal phase HPLC separation of 2AB glycans are presented in Figure 11 and Figure 12. Structural identification of the permethylated N-glycans were assigned by searching for the same ion in the

MALDI structural databases and reference spectra of mouse brain glycans provided by the Consortium for Functional Glycomics (407). A majority of the permethylated glycans extracted from the tissues were the same masses as the glycans listed in the Consortium for Functional Glycomics mouse brain tissue spectral libraries.

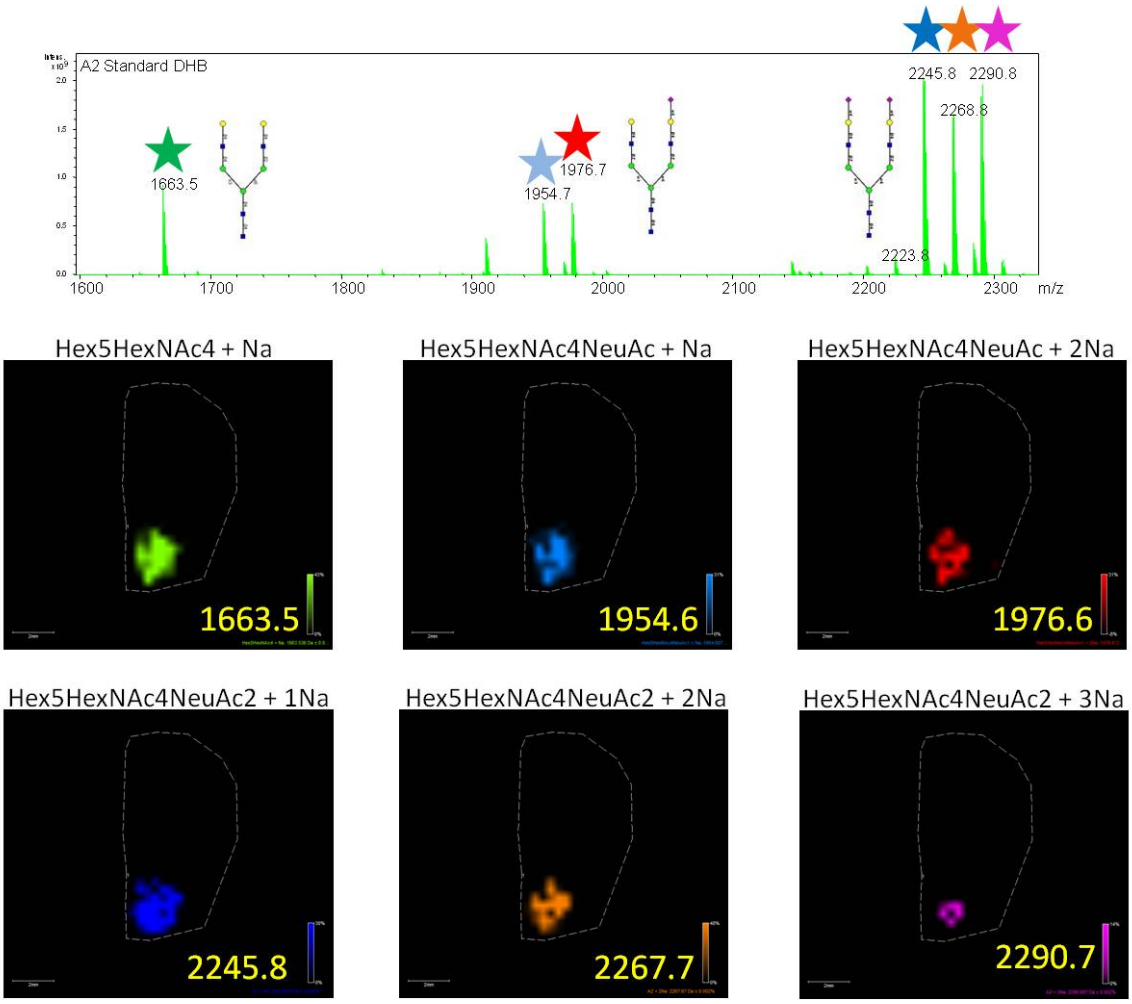


Figure 10. Analysis of A2 Standard on Tissue Sections. An A2 Standard (Hex5HexNAc4NeuAc2; 0.1 µg) was spotted directly on a ground steel plate (a.), or 0.1 µg/µL standard was spotted onto a mouse brain tissue slice prior to matrix application (b-g). The MALDI mass spectrum demonstrates the spectral complexity of non-derivatized native sialic acid N-glycans due to loss of sialic acid residues and different numbers of sodium adducts. Representative images correspond to the ions detected in the mass spectrum localized to the tissue area where they were spotted (b-g). The glycan composition and number of sodium adducts per analyte are indicated above each image panel. The colored stars in the spectra correspond to the colors in the image panel.

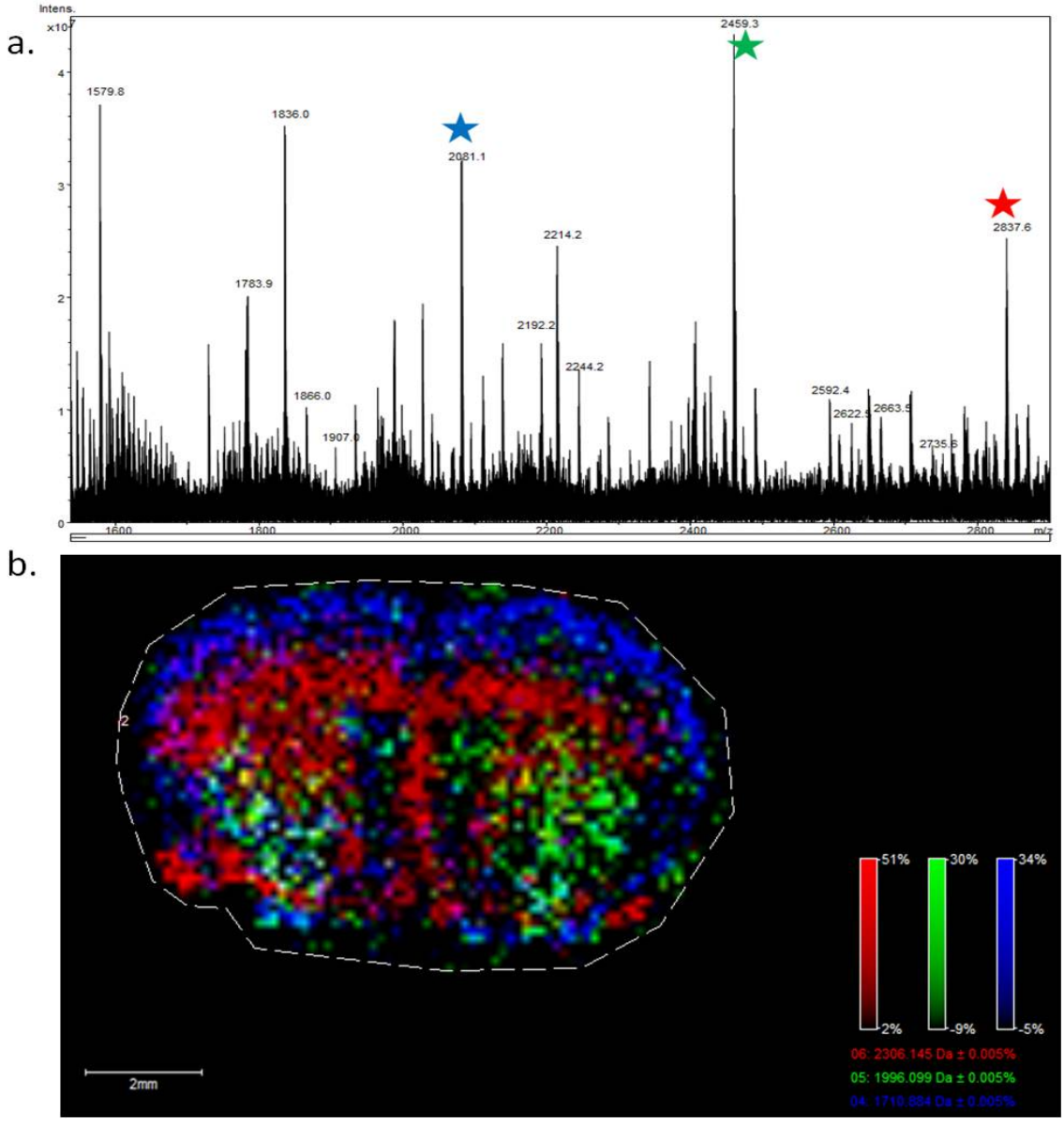
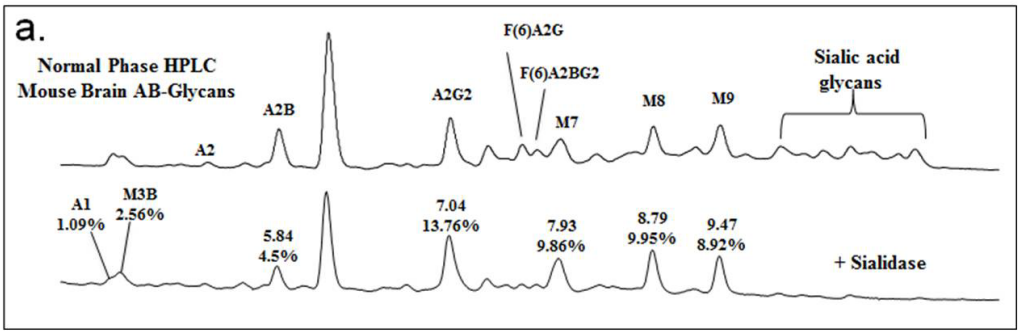


Figure 11. Linkage of Permethylated Glycan Masses with Native Tissue Glycan Ion Distribution. The N-glycan imaging method was applied to a mouse brain for tissue imaging and an off tissue permethylation was performed for glycan verification. (a.) Permethylated glycans Hex5dHex3HexNAc5 + Na (m/z = 2837.6), Hex4dHex2HexNAc5 + Na (m/z = 2459.3), and Hex3dHex1HexNAc5 +2Na (m/z = 2081.1) were identified and the corresponding glycans displayed specific localization in the tissue image (b). M/z = 2306.1 corresponds to Hex5dHex3HexNAc5 + 2Na (red), m/z = 1996.1 corresponds to Hex4dHex2HexNAc5 + Na (green), and m/z = 1710.9 corresponds to Hex3dHex1HexNAc5 + 2Na (blue). A raster distance of 125 microns was used per MALDI laser acquisition across the tissues.



A1 – Hex3HexNAc3
A2 – Hex3HexNAc4
A2G2 – Hex5HexNAc4
F(6)A2BG2- Hex5dHex1HexNAc5
M8 - Hex8HexNAc2
M3B – Hex5HexNAc3
A2B – Hex3HexNAc5
F(6)A2G – Hex4dHex1HexNAc4
M7 - Hex7HexNAc2
M9 - Hex9HexNAc2

Figure 12. HPLC Analysis of 2-AB Labeled N-Glycans. Mouse brain tissues (10 um slices) were digested on-slide with PNGaseF, extracted with water then concentrated for derivation with 2-aminobenzoate (2AB) and separated by normal phase HPLC. Sialidase digestion was done in solution with the 2AB modified glycans, prior to HPLC separation. The structure key for indicated normal phase peaks in panel below the chromatogram.

Figure 13. Mouse Brain N-Glycan Imaging Panel. 28 N-glycans were observed in both the MALDI-IMS dataset and the permethylation spectra. These ions vary with respect to brain histology**.**

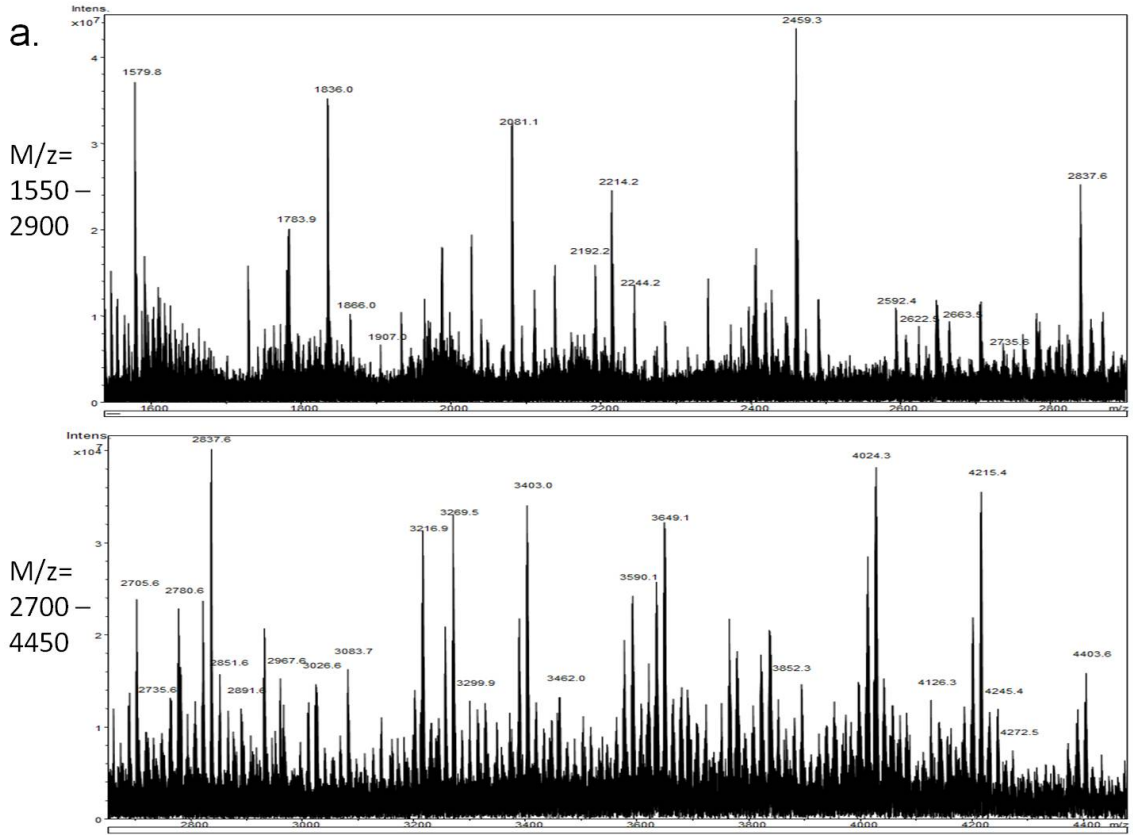


Figure 14. N-Glycan Permethylation Spectra. Permethylation analysis of the same glycan extract as done with the HPLC analysis, with MALDI spectra split into two overlapping mass ranges: (top panel) m/z = 1550-2900; (bottom panel) m/z = 2700-4550.

These permethylated masses and structures were used to correlate back to the expected native mass/structure of the glycan. As an example, an overlay of Hex5dHex3HexNAc5 + Na, Hex4dHex2HexNAc5 + Na, and Hex3dHex1HexNAc5 + Na demonstrates the spatial pattern of expression for these specific glycans by MALDI-IMS linked to the corresponding extracted permethylated peaks (Figure 11). A reference panel of 28 glycan structures was generated that lists structural composition (Figure 13), expected and observed masses of the permethylated glycans (Figure 14), and the masses of their corresponding native glycan masses for cross-referencing permethylated glycan spectra and native tissue glycan masses. Additionally, the majority of the identities of the off-tissue native glycan species reacted with 2AB and separated by normal phase HPLC are also common to those listed in Figure 13. The majority of the detected masses were within 0.5 mass units of the predicted masses. The few discrepancies of more than 0.5 mass units generally occurred with native glycans of larger mass (>2500) and were associated with complex overlapping ion signals in those specific tissue regions.

Many larger glycans are branch-chained with terminal sialic acid residues, thus a comparison was done for detection of on-tissue glycans released by PNGaseF digestion alone versus a combination of PNGaseF and sialidase. If larger sialylated structures were present, then sialidase digestion would result in new glycan peaks detected in the spectra relative to PNGaseF alone. As shown in Figure 15, the on-tissue analysis indicated an increase in the levels of larger mass native glycan ions (m/z > 2500) detected in the double digest of PNGaseF and sialidase (Figure 15a–c). These structures are consistent with the sialidase products of glycans that contained tri- and tetra branched chains with multiple N-acetylglucosamine-galactose (LacNAc) dimers (408). A representative overlay image of three of these structures detected only in the combined sialidase/PNGaseF digest is shown in Figure 16d. Sialidase digestion of 2AB modified glycans also demonstrated shifts in the detection of smaller glycan species following normal phase HPLC separation (Figure 12).

The methods developed for mouse brain tissues were applied to a frozen slice of human kidney containing both medulla and cortex regions. As shown in Figure 17, representative N-glycans released by on-tissue PNGaseF digestion can be detected that are specific to either cortex or medulla regions. Compared to the mouse brains analyzed, the total number of glycan species was slightly reduced in the human kidney sections. Beyond these kidney tissues and mouse brain tissues, we have found it is necessary to optimize the length of incubation and amount of PNGaseF digestion conditions for each tissue type being tested. Overall, these baseline conditions described for mouse brains have worked as useful starting points for any new tissue being analyzed.

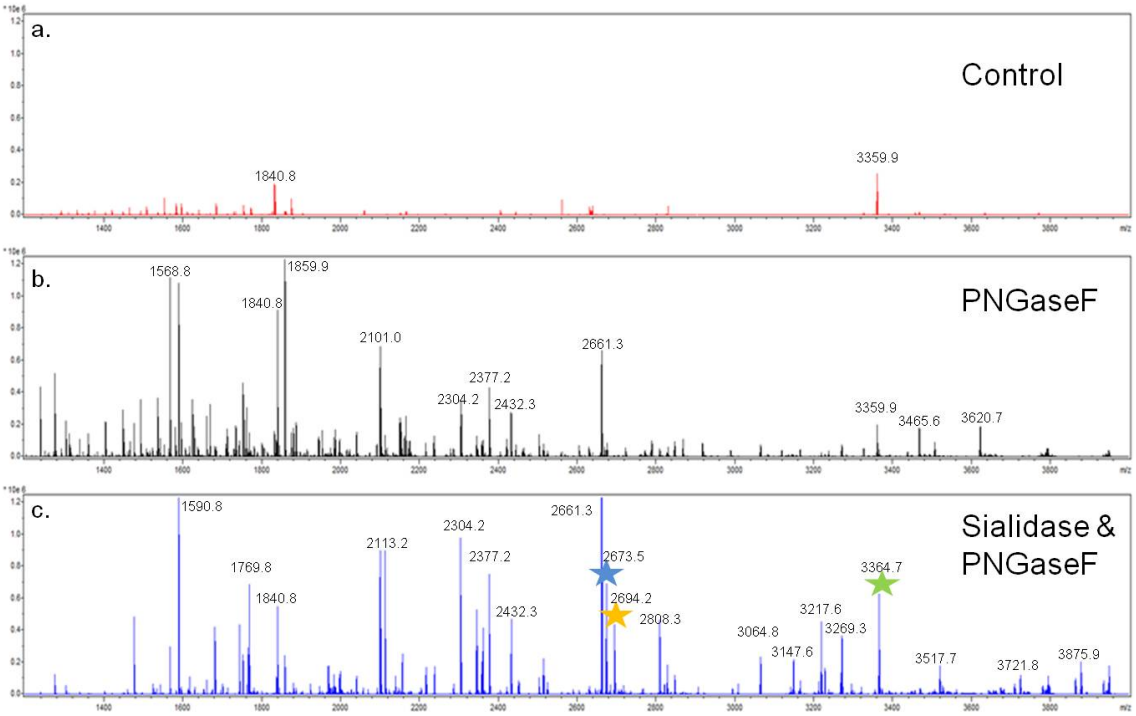


Figure 15. Mass Spectra of N-Glycans Released with PNGaseF and Sialidase. Three serial sections from a mouse brain were prepared and either (a) untreated, (b) digestion with PNGaseF, or (c) a combination of PNGaseF and Sialidase S. Recombinant PNGaseF (20 mU) alone or in combination with 100 mU Sialidase S (Prozyme) were applied using a Bruker ImagePrep prior to 2 hr incubation at 37° C.

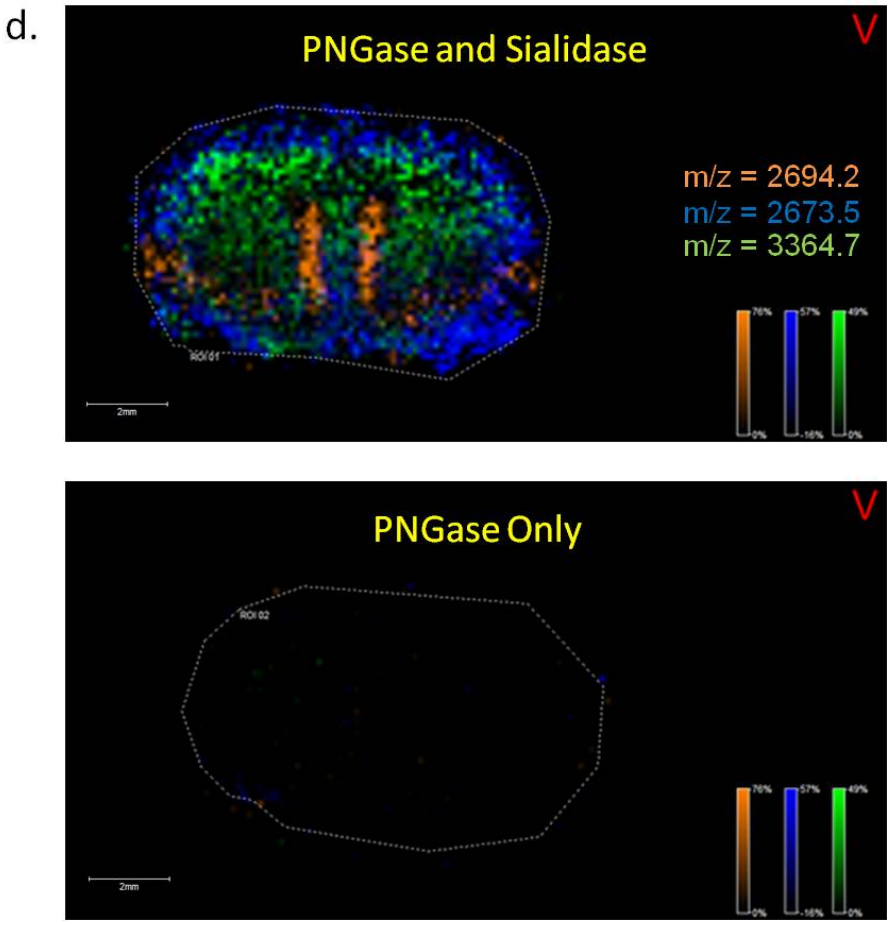


Figure 16. Novel Glycans Observed After Sialidase Release. MALDI-MSI results of the average spectra shown in Figure 15, where a representative overlay image of the unique glycans detected with the sialidase and PNGaseF combination is shown for m/z values of 2673.5 (blue), 2694.2 (orange) and 3364.7 (green). A raster distance of 125 microns was used per MALDI laser acquisition across the tissues.

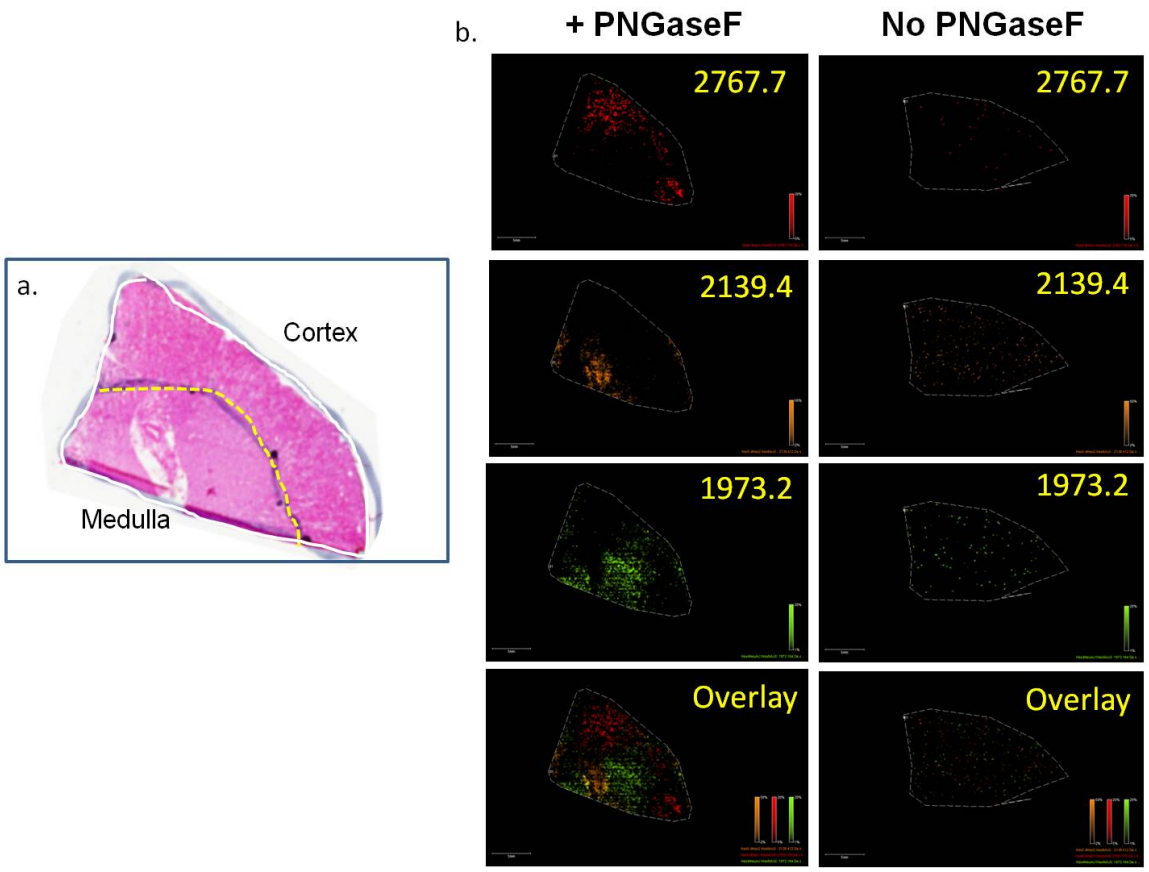


Figure 17. N-Glycan MALDI-IMS of a Normal Human Kidney Tissue Section. Three serial sections from a human kidney were prepared and either untreated (data not shown), digested with PNGaseF, or analyzed by H&E staining. Cortex and medulla regions were confirmed by a pathologist. Recombinant PNGaseF (20 mU) was applied using a Bruker ImagePrep followed by a 2 hr incubation at 37° C. Following MALDI-MSI, m/z values that are specific to PNGaseF treated tissue are observed. Shown are the results for an m/z value of 2767.7 (a) localized in the cortex, while the m/z values of 2139.4 (b) and 1973.2 (c) are localized primarily in the medulla when compared to the H&E stain (d). An overlay of the three m/z values is provided (e). A raster distance of 125 microns was used per MALDI laser acquisition across the tissues.

3.2.5 Discussion

We describe a new mass spectrometry-based glycan imaging approach in combination with on-tissue PNGaseF digestion to spatially profile released N-linked glycans in their local microenvironment. The method has been designed to facilitate the tissue analysis of the cell surface glycans in frozen tissues while maintaining histopathology compatible preparation workflows. While MALDI-IMS has primarily been utilized to spatially profile proteins, lipids, drug and small molecule metabolites in tissues, it has not been previously applied to N-linked glycan analysis. Coronal slices of mouse brain tissues were ideal for method development because of the robust glycan signal and the existence of an N-glycan catalogue, provided by the Consortium for Functional Glycomics, that documents known N-glycans present in mouse brain tissue (407). Therefore the off-tissue extraction of glycans from PNGaseF digested brain slices for permethylation and mass spectrometry profiling or normal phase HPLC analyses could be readily compared to on-tissue glycan species, in particular for assigning glycan structure and composition. The Consortium of Functional Glycomics reference spectra library of permethylated N-glycan species from mouse brains as detected by MALDI lists over 40 identified structures. We estimate that 30 or more N- glycan species can be detected on-tissue in mouse brains by our current MALDI-IMS workflows, including the 28 structures listed in the panels of Figure 13. This number varies depending on the location of the tissue cut within the brain and the number of sub-regions present. MALDI-IMS detection of brain substructure regions of protein, small molecule metabolites and lipid distributions in brain substructure regions is well documented, so detection of specific N-linked glycan isoform distribution is not surprising, and represents to our knowledge the first demonstration of this (396,409–411). The specific distribution of N-glycans in the medulla and cortex regions of the human kidney tissue (Figure 17) also demonstrates sub-organ structure specificity of glycan species expression.

MALDI mass spectrometry is a common detection method for analysis of complex carbohydrates, including routine analysis of N-linked glycans released from glycoproteins by PNGaseF. From a practical perspective, our approach used a recombinant PNGaseF enzyme engineered for increased stability and activity. In addition, the amount of commercially available preparations of PNGaseF that would be required to perform the described experiments would not be economically feasible for most laboratories. The recombinant enzyme is also not formulated with detergents and additives that are common components of commercial enzyme, many of which can interfere with MALDI ionization. However, glycan analysis by MALDI is notoriously complicated by problems with glycan dissociation and sodium adduct additions. It is common to observe the loss of sialic acid when native glycans are analyzed by MALDI MS, hence the need for modifications like permethylation or 2-aminobenzoate that stabilize and help ionize sialic acid containing glycans (247). Most common chemical modification methods require organic solvents and the need to separate reaction products, thus these are not feasible for on-tissue derivatization and retention of spatial distributions. The MALDI-IMS profiling of the A2 glycan standard on-tissue was similarly affected, with the number of Na ions and loss of sialic acid being a factor for an effective prediction of structure. In addition to those glycans listed in Figure 13 that we have cross-referenced for structure identification, there are multiple PNGaseF-released ions detected that have not been correlated to reported structures. Many of these are larger species m/z > 3500, and like those glycan ions assigned structures, all will ultimately require confirmatory structural characterization involving tandem MS identifications off tissue (412,413). Use of MALDI alone may be difficult to structurally identify very high mass poly-lactosamino-glycan species on-tissue, but some of these higher mass glycans are potentially profilable by MALDI-IMS (408). Differential exoglycosidase digestion strategies more amenable to native glycan identification on-tissue and preservation of spatial distribution, analogous to that described for sialidase (Figure 15), are ongoing. Preliminary experiments in negative ion mode using other matrices have

yielded promising results, and we expect this could be a partial solution for more effective detection of native sialylated glycan species.

Overall, we present a functional MALDI-IMS workflow for on-tissue analysis and distribution of N-glycan species, and there remains room for continued improvement of the assay and analysis workflow. We only report the use of one type of MALDI matrix, data was only collected in positive ion mode, and only fresh frozen tissues were used. Except for those glycans directly compared with reference standards, the structural assignments are correlative to existing databases and subject to known limitations of MALDI analysis. Further verification of these assigned structures by tandem mass spectrometry approaches is ongoing. Evaluation of different chemical matrices for either positive or negative ion mode analysis are ongoing, as well as refinement of methods to profile N-glycans in formalin- fixed paraffin-embedded (FFPE) tissues. As MALDI-IMS has been used to effectively profile proteins in FFPE tissues, we expect that these methods should be adaptable for glycan analysis in FFPE tissues (144,414). We believe this capability of profiling N-glycans spatially on tissue will have significant application to tissues representing many types of human diseases, especially the changes in glycosylation associated with oncogenesis. It is well documented that malignant transformation and cancer progression result in fundamental changes in the glycosylation patterns of cell surface and secreted glycoproteins, and the majority of current FDA-approved tumor markers are glycoproteins or glycan antigens (209,412,415,416). Therefore, development of a molecular assay like the MALDI-IMS approach we describe could directly profile these pathologically relevant carbohydrates in their local microenvironment and accelerate approaches to aid in the diagnosis, prognosis, staging, and treatment of many types of cancers. Coupled with many new improved high resolution mass spectrometry options for intact glycopeptides and small molecule carbohydrate binding probes, there is potential to link the tissue distribution of glycans from MALDI-IMS to the glycoproteins carrying them (320,321,417,418).

Acknowledgments

This work was supported by grants from the National Institutes of Health/National Cancer Institute grants R21CA137704 and R01CA135087, and the state of South Carolina SmartState Endowed Research program to R.R.D; This work was supported by grants R01 CA120206 and U01 CA168856 from the National Cancer Institute (NCI), the Hepatitis B Foundation, and an appropriation from The Commonwealth of Pennsylvania to A.M. This work was supported by NIH/NCI grants R01CA104505 and R01CA104505-05S1, a generous gift from the David & Lois Stulberg Endowed Fund for Kidney Cancer Research, Kidney Cancer Research at Mayo Clinic in Florida and The Fraternal Order of Eagles, State of Florida to J.A.C.

3.3 Adaptation of MALDI-IMS Workflow for FFPE Tissue Blocks

Manuscript published in *PLoS One*, September 2014 9(9): e106255. TWP performed all mass spectrometry based expriments, as well as constructed the figures and text of the manuscript. RRD contributed in the editing and design of the manuscript. BAN designed the TMA data processing strategy to account for MALDI-IMS data. YS served as the pathologist to identify specific regions of the tissue.

MALDI Imaging Mass Spectrometry Profiling of N-Glycans in Formalin-Fixed Paraffin Embedded Clinical Tissue Blocks and Tissue Microarrays.

Thomas W. Powers[1], Benjamin A. Neely[1], Yuan Shao[2], Huiyuan Tang[5], Dean A. Troyer[3], Anand S. Mehta[4], Brian B. Haab[5] and Richard R. Drake*[1,2]

1. Department of Cell and Molecular Pharmacology and Experimental Therapeutics and MUSC Proteomics Center, Medical University of South Carolina, Charleston, SC, USA

2. Hollings Cancer Center Biorepository and Tissue Analysis Resource, Medical University of South Carolina, Charleston, SC, USA

3. Departments of Pathology and Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia, USA.

4. Drexel University College of Medicine, Department of Microbiology and Immunology and Drexel Institute for Biotechnology and Virology, Doylestown, PA USA

5. Laboratory of Cancer Immunodiagnostics, Van Andel Research Institute, Grand Rapids, MI, USA

* To whom correspondence should be addressed: Richard R. Drake, 173 Ashley Avenue, Charleston, SC, 29425; (tel) 843 792 4505: draker@musc.edu

Key Words: N-linked glycosylation, formalin-fixed paraffin-embedded tissue, tissue microarray, MALDI imaging mass spectrometry, glycoprotein

3.3.1 Abstract

A recently developed matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS) method to spatially profile the location and distribution of multiple N-linked glycan species in frozen tissues has been extended and improved for the direct analysis of glycans in clinically derived formalin-fixed paraffin-embedded (FFPE) tissues. Formalin-fixed tissues from normal mouse kidney, human pancreatic and prostate cancers, and a human hepatocellular carcinoma tissue microarray were processed by antigen retrieval followed by on-tissue digestion with peptide N-glycosidase F. The released N-glycans were detected by MALDI-IMS analysis, and the structural composition of a subset of glycans could be verified directly by on-tissue collision-induced fragmentation. Other structural assignments were confirmed by off-tissue permethylation analysis combined with multiple database comparisons. Imaging of mouse kidney tissue sections demonstrates specific tissue distributions of major cellular N-linked glycoforms in the cortex and medulla. Differential tissue distribution of N-linked glycoforms was also observed in the other tissue types. The efficacy of using MALDI-IMS glycan profiling to distinguish tumor from non-tumor tissues in a tumor microarray format is also demonstrated. This MALDI-IMS workflow has the potential to be applied to any FFPE tissue block or tissue microarray to enable higher throughput analysis of the global changes in N-glycosylation associated with cancers.

3.3.2 Introduction

Tissues obtained from surgeries or diagnostic procedures are most commonly preserved in formalin-fixed paraffin-embedded (FFPE) tissue blocks. These tissues are fixed in formalin and processed as paraffin-embedded tissue blocks. The embedding process preserves the cellular morphology and allows tissues to be stored at room temperature, causing FFPE fixation to be used by many tissue banks and biorepositories (419,420). For cancer biomarker discovery, FFPE

tissues are particularly attractive because they are archived for years and are much more widely available than cryopreserved tissue. When combined with clinical outcomes, FFPE tissues are a rich source of samples for biomarker discovery and validation in retrospective studies. While the fixation method has many benefits, the formalin treatment results in the formation of methylene bridges between the amino acids of the proteins, complicating further analysis by mass spectrometry. There has been continued progress in improving extraction methods of trypsin digested peptides from FFPE tissues in recent years, in parallel with improved high resolution sequencing analysis of peptides by mass spectrometry (387,421). Incorporation of multiple FFPE tumor tissue cores in a tissue microarray (TMA) format also has proven to be effective for immunohistochemistry analysis of potential biomarker candidates, and TMAs are increasingly being used for validation of alterations in protein expression associated with emerging genetic mutation phenotypes and transcriptional profiling studies (422–424). The main advantages of experiments performed with TMAs are the ability to include multiple cores from the same subject tumors, improved sample throughput, statistical relevance and multiplexed analysis of diverse molecular targets (422,425). Thus, it is possible to place up to 100 samples with duplicates and controls on a single slide. When correlated with associated clinical outcomes, this provides a powerful method for biomarker discovery and validation while minimizing reagent use and assuring that each core in the TMA is treated under identical conditions.

It is well documented that malignant transformation and cancer progression result in fundamental changes in the glycosylation patterns of cell surface and secreted glycoproteins (209,415,416). Glycosylation of proteins are post-translational modifications most commonly involving either N-linked addition to asparagine residues or O-linked additions to serine or threonine residues. Current approaches to evaluate glycosylation changes generally involve bulk extraction of glycans and glycoproteins from tumor tissues for analysis by mass spectrometry or antibody array platforms, however, this disrupts tissue architecture and distribution of the

analytes. Broad affinity carbohydrate binding lectins and a small number of glycan antigen antibodies can be used to target glycan structural classes in tissues, but not individual glycan species. Additionally, these detection methods for global alterations in glycosylation requires staining on many adjacent tissue sections, making large scale assessments on many samples difficult, expensive and time consuming. There are only a few reported studies examining glycosylation related changes of proteins or glycolipids in FFPE cancer tissues, and these focus primarily on determining the levels of the protein carriers or glycosyltransferases through immunostaining (426–428).

One potential approach to assess glycan changes in tissues is matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS). This technique has been used to directly profile multiple protein, lipid and drug metabolite in tissue, generating molecular maps of the relative abundance and spatial distribution of individual analytes linked to tissue histopathology (138,143,395–399). MALDI-IMS analysis of peptides following trypsin digestion of FFPE TMAs have also been reported (144,429,430). Recently, our group reported a MALDI-IMS method workflow to directly profile N-linked glycan species in fresh/frozen tissues (431). Adapting this method for the analysis of N-glycans in FFPE tissues would serve to extend the application of the technique to larger retrospective sample sets and TMAs.

In this report, we describe the application of MALDI-IMS glycan imaging to various formalin-fixed tissues. Formalin-fixed mouse kidney tissues were used to optimize antigen retrieval, PNGaseF digestion and glycan detection conditions for MALDI-IMS. This was followed by N-glycan analysis of clinical FFPE tissue blocks from prostate and pancreatic cancers, as well as a commercial tissue microarray of hepatocellular carcinoma (HCC). Glycan identity was confirmed by on-tissue collision-induced dissociation (CID) and off-tissue permethylation analysis. An optimized MALDI-IMS workflow is presented that allows routine simultaneous analysis of 30 or more glycans per FFPE tissue, including TMA formats. The

approach is amenable to any FFPE tissue, and represents an additional molecular correlate assay for use with the TMA format. Furthermore, depending on the construction of the TMA and targeted tumor type, the approach has the potential to identify novel glycan biomarker panels for cancer detection and prognosis. To our knowledge, this represents the first instance of using MALDI-IMS to profile N-glycans in FFPE tissue blocks or TMAs.

3.3.3 Materials and Methods

3.3.3.1 Materials

The glycan standard NA2 was obtained from ProZyme (Hayward, CA). Trifluoroacetic acid, sodium hydroxide, dimethyl sulfoxide, iodomethane and α-cyano-4-hydroxycinnamic acid (CHCA) were obtained from Sigma-Aldrich (St. Louis, MO). HPLC grade methanol, ethanol, acetonitrile, xylene and water were obtained from Fisher Scientific (Pittsburgh, PA). ITO slides were purchased from Bruker Daltonics (Billerica, MA) and Tissue Tack microscope slides were purchased from Polysciences, Inc (Warrington, PA). Citraconic anhydride for antigen retrieval was from Thermo Scientific (Bellefonte, PA). Recombinant Peptide N-Glycosidase F (PNGaseF) from *Flavobacterium meningosepticum* was expressed and purified as previously described (431).

3.3.3.2 FFPE Tissues and TMAs

All human tissues used were de-identified and determined to be not human research classifications by the respective Institutional Review Boards at MUSC and Van Andel. Mouse kidneys were excised from euthanized C57BL/6 mice and immediately placed in 10% formalin prior to processing for routine histology and paraffin embedding. Mice were housed in an Institutional Animal Care and Use Committee-approved small animal facility at MUSC, and tissues obtained were harvested as part of approved projects unrelated to glycan tissue imaging. A liver TMA was purchased from BioChain consisting of 16 cases of liver cancer in duplicates, and one adjacent non-tumor tissue for each case. Tissues were from 14 male and two female patients

with an average age of 47.5 with a range of 33 to 68 years old, with additional information provided in Supplementary Table 1. A de-identified prostate tumor FFPE block, stored for 10 years representing a Gleason grade 6 (3+3)/stage T2c adenocarcinoma from a 62 year old Caucasian male, was obtained from the Hollings Cancer Center Biorepository at the Medical University of South Carolina. A pathologist confirmed the presence of approximately 10% prostate cancer gland content in the sample. A de-identified large-cell undifferentiated pancreatic carcinoma FFPE tissue section with low CA19-9 staining was obtained from the Van Andel Institute Biospecimen Repository. For each section analyzed, histological analysis and staining with hematoxylin and eosin (H&E) were performed.

3.3.3.3 Washes for Deparaffinization and Rehydration

Tissue and TMA blocks were sectioned at 5μm and mounted on positively charged glass slides measuring 25 x 75mm, compatible with the Bruker slide adaptor plate. The slides were heated at 60°C for 1hr. After cooling, tissue sections were deparaffinized by washing twice in xylene (3 minutes each). Tissue sections were then rehydrated by submerging the slide twice in 100% ethanol (1 minute each), once in 95% ethanol (one minute), once in 70% ethanol (one minute), and twice in water (3 minutes each). Following the wash, the slide was transferred to a coplin jar containing the citraconic anhydride buffer for antigen retrieval and the jar was placed in a vegetable steamer for 25 minutes. Citraconic anhydride (Thermo) buffer was prepared by adding 25μL citraconic anhydride in 50mL water, and adjusted to pH 3 with HCl. After allowing the buffer to cool, the buffer was exchanged with water five times by pouring out ½ of the buffer and replacing with water, prior to replacing completely with water on the last time. The slide was then desiccated prior to enzymatic digestion. Tris buffer pH 9-10 was also effective, but citraconic anhydride buffer was used for all experiments in this study.

3.3.3.4 N-Glycan MALDI-IMS

An ImagePrep spray station (Bruker Daltonics) was used to coat the slide with a 0.2mL aqueous solution of PNGaseF (20µg total/slide) as previously described (431). Adjacent control tissue slices lacking PNGaseF were generated by covering them with a glass slide during the spraying process. Following application of PNGaseF, slides were incubated at 37°C for 2hr in a humidified chamber, then dried in a desiccator prior to matrix application. The matrix, α-Cyano-4-hydroxycinnamic acid matrix (0.021g CHCA in 3mL 50% acetonitrile/50% water and 12µL 25%TFA) was applied using the ImagePrep sprayer. Released glycan ions were detected using a SolariX dual source 7T FTICR mass spectrometer (Bruker Daltonics) (m/z = 690-5000 m/z) with a SmartBeam II laser operating at 1000 Hz, a laser spot size of 25µm. Images of differentially expressed glycans were generated to view the expression pattern of each analyte of interest using FlexImaging 4.0 software (Bruker Daltonics). Following MS analysis, data was loaded into FlexImaging Software focusing on the range m/z = 1000-4000 and reduced to 0.95 ICR Reduction Noise Threshold. Observed glycans were searched against the glycan database provided by the Consortium for Functional Glycomics (www.functionalglycomics.org). Glycan structures were generated by Glycoworkbench and represent putative structures determined by combinations of accurate m/z, CID fragmentation patterns and glycan database structures (432).

3.3.3.5 Permethylation of Tissue Extracted N-glycans:

PNGaseF sprayed mouse kidney tissue slides were incubated for 2hr at 37°C; 50µL water was applied on top of the tissue and incubated for 20 minutes to extract the released native N-glycans. The water was removed from the tissue, and then concentrated under vacuum by centrifugation. Permethylation was performed as described (431), and glycans analyzed by MALDI. Masses detected in the permethylation experiments were searched against the permethylated glycan database provided by the Consortium for Functional Glycomics (407).

3.3.3.6 Collision-Induced Dissociation of N-linked Glycans

Glycan standards were spotted on a stainless steel MALDI plate using CHCA matrix and desiccated to yield a homogenous layer. Tissues were prepared as previously described for MALDI imaging of FFPE tissues. 10 spectra of 1000 laser shots with a laser frequency of 1000 Hz were averaged for each spectra provided. The collision energy varied between 60-70V.

3.3.3.7 TMA Statistics

Mass spectra from TMA tissue Regions of Interest (ROIs) representing each tissue core were exported directly from FlexImaging and analyzed using an in-house workflow. The peak lists were first deconvoluted followed by calculating the mean peak intensity of points in each ROI, resulting in a monoisotopic peak list corresponding to signal intensity in each region. Comparison of tumor versus non-tumor was accomplished with a Wilcoxon rank sum test. Individual peaks were also evaluated to discriminate between tumor and non-tumor using receiver operator characteristic curves.

3.3.4 Results

3.3.4.1 Analysis of Formalin-Fixed Mouse Kidneys and Human Cancer Tissues.

Mouse kidney tissues were fixed in formalin and used as an initial model system to develop MALDI-IMS glycan imaging workflows for FFPE tissues. These tissues were chosen due to the availability of reference glycan structures and spectra, and previous MALDI-IMS glycan imaging data from our laboratory for fresh/frozen tissue analysis (407,431). A summary workflow schematic is provided (Figure 18). Tissues were cut at 5 microns, deparaffinized and rehydrated in sequential xylene/ethanol/water rinses, followed by antigen retrieval in citraconic anhydride pH 3. The rehydrated tissues were sprayed with PNGaseF, incubated for glycan release, sprayed with CHCA matrix, and then analyzed by MALDI-IMS. While all data shown herein uses CHCA, 2,5-dihydroxybenzoic acid (DHB) matrix could also be used successfully for

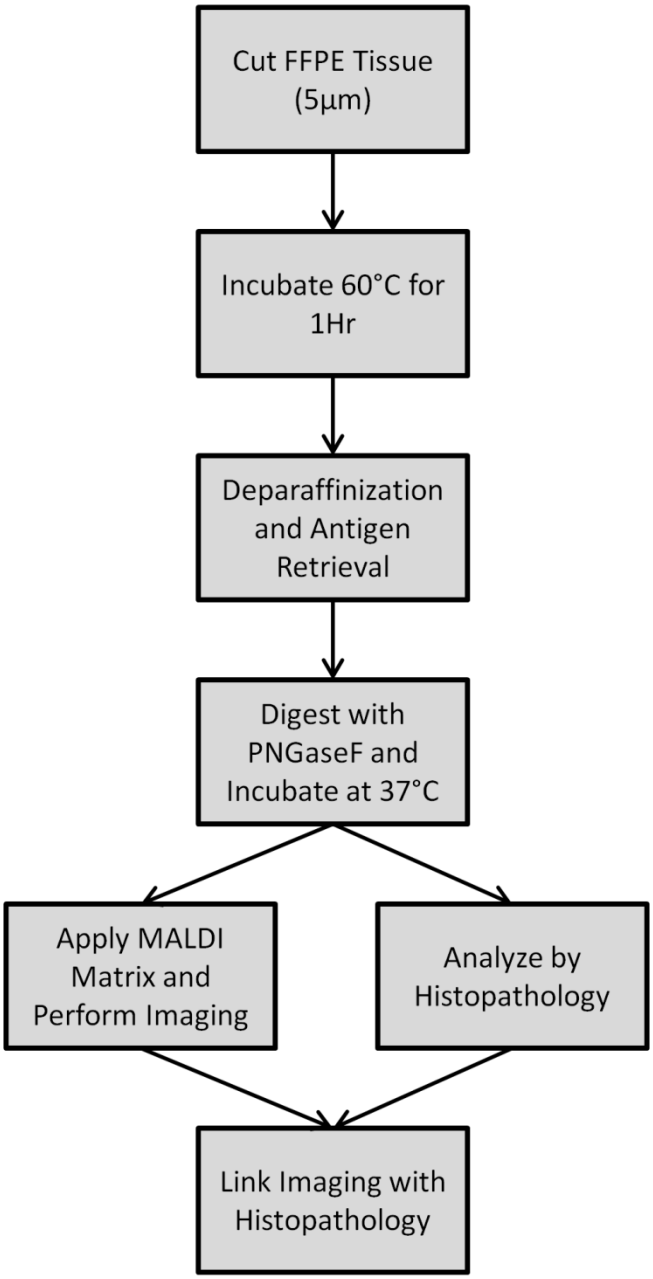


Figure 18. MALDI-IMS Workflow for Imaging N-Glycans on FFPE Tissues

N-glycan imaging of FFPE tissues. As shown in Figure 19, there were multiple ions detectable only in the tissue incubated with PNGaseF that were not present in the control tissue with no PNGaseF application. Different glycans were distributed across the cortex or medulla regions. For example, a Hex4dHex2HexNAc5 ion (m/z = 1996.74) is present in the cortex and medulla (Figure 19c), while a Hex5dHex2HexNAc5 glycan (m/z = 2158.76) is more specific to the cortex (Figure 19d). An overlay of the MALDI-IMS images for these two ions from the PNGaseF treated sections (Figure 19e) and the control tissue (Figure 19f) demonstrates that these two ions are released by PNGaseF. In control kidney tissues that were only sprayed with aqueous PNGaseF solution lacking enzyme, or tissue slices that were not processed by antigen retrieval plus and minus PNGaseF digestion, only matrix ions or paraffin/formalin polymer were detected (data not shown). A summary glycan image panel of 28 glycan ions detected in these kidneys, sodium adducts and observed/expected m/z values is provided in Figure 20. Additionally, N-glycans were extracted from the tissue following on-tissue PNGaseF digestion, permethylated and analyzed by MALDI. A representative spectra from this analysis is provided in Figure 21. These permethylated values were also compared with MALDI reference spectra for mouse kidney glycans from the Consortium for Functional Glycomics. The imaged glycan ions were correlated to the reference spectra glycans, illustrated in Figure 22, and could be matched to all 28 glycan species highlighted in the reference spectra.

We next assessed whether the method was compatible with two representative archived pathology FFPE tissue blocks, one for pancreatic cancer and one for prostate cancer. A section of human pancreatic cancer tissue of complex histology was processed, incubated with PNGaseF and glycans detected by MALDI-IMS (Figure 23). Different N-glycans were detected that could distinguish between non-tumor, tumor, tumor necrotic and fibroconnective tissue regions. A representative glycan image overlay of four m/z values that correspond to the sodium adducts of potential N-glycan species is shown in Figure 23a, each representing a specific region of the

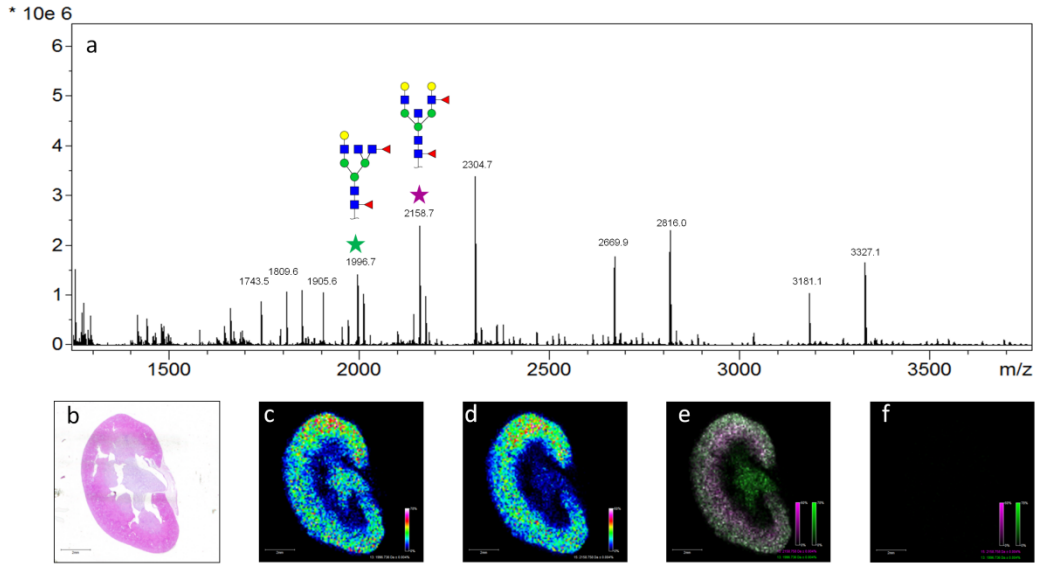


Figure 19. MALDI-IMS of N-Glycans on a Mouse Kidney Tissue Section. Two mouse kidneys were sliced at 5 µm prior to proceeding with the MALDI-IMS workflow. One tissue was covered with a glass slide during PNGaseF application to serve as an undigested control tissue. An average annotated spectra from the tissue that received PNGaseF application is provided (a). Tissue regions were assessed by H&E stain (b). The labeled peaks correspond to native N-glycans that have been reported for the mouse kidney on the Consortium for Functional Glycomics mouse kidney database. Two of these ions were selected and their tissue localization was assessed. Hex4dHex2HexNAc5 at m/z=1996.7 (c) is located in the cortex and medulla while Hex5dHex2HexNAc5 m/z=2158.7 (d) is more abundant in the cortex of the mouse kidney. An overlay image of these two masses is also shown (e), as well as the corresponding image from untreated PNGaseF control tissues (f).

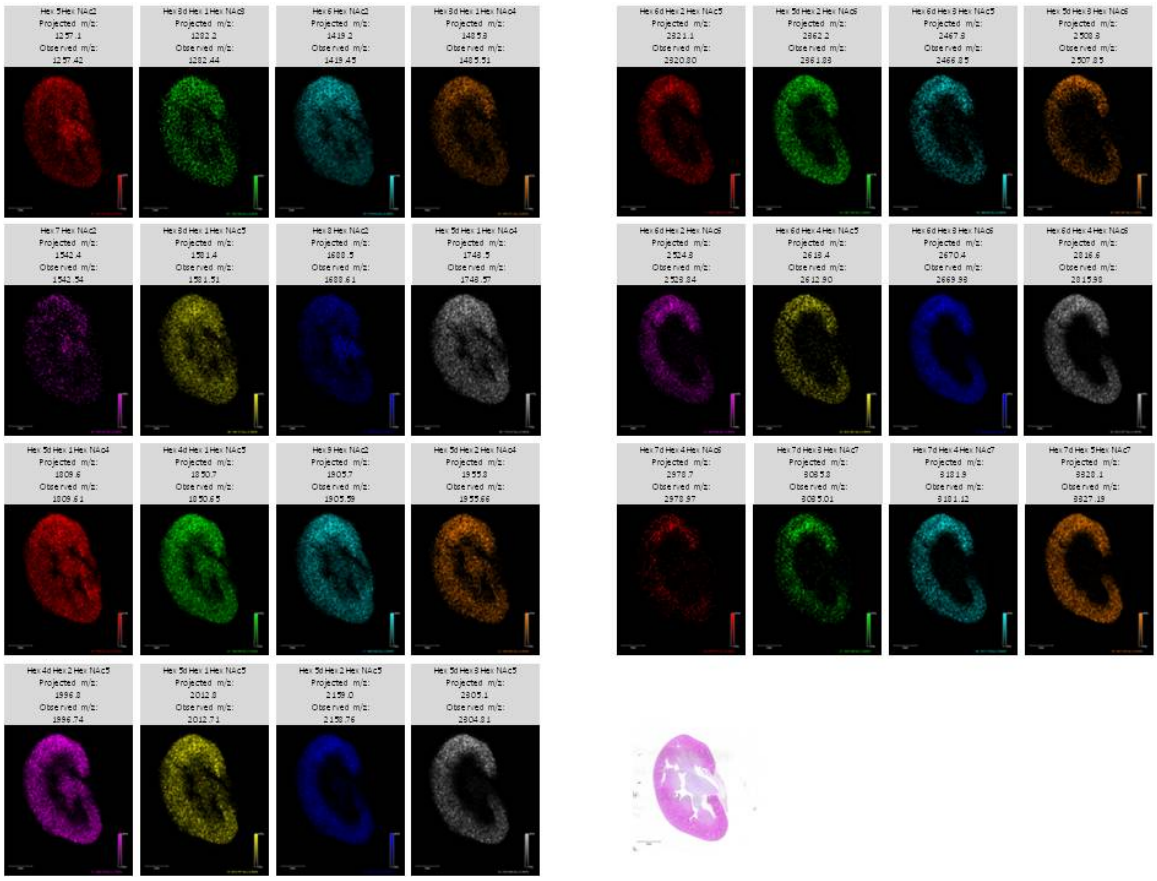


Figure 20. N-Glycan Panel of FFPE Kidney Sections. Ions detected in the kidney with enzyme application were compared to the control tissue. Ions that were only observed in the tissue following PNGaseF application were compared to the glycans found in the mouse kidney database on the Consortium for Functional Glycomics. The panel provides the glycan species, the projected mass for the sodium adduct, and our observed mass for the sodium adduct.

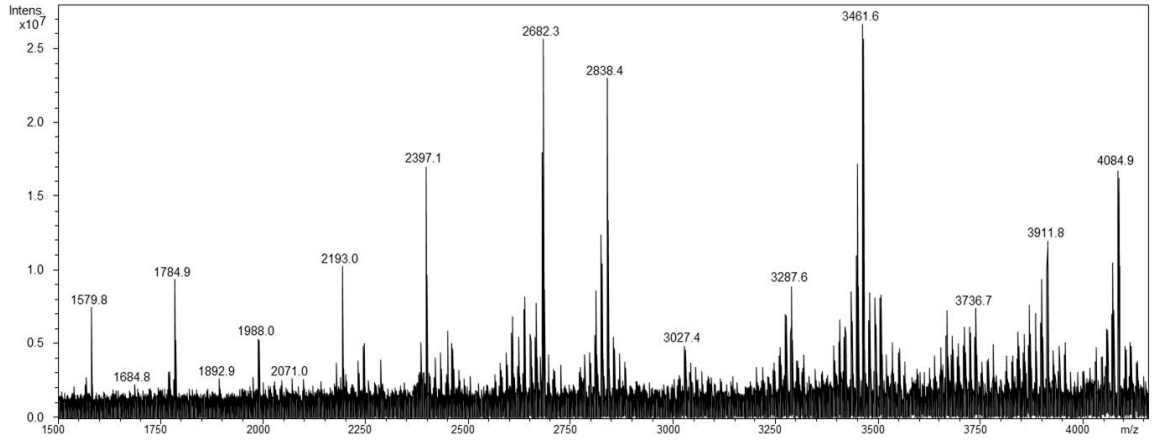
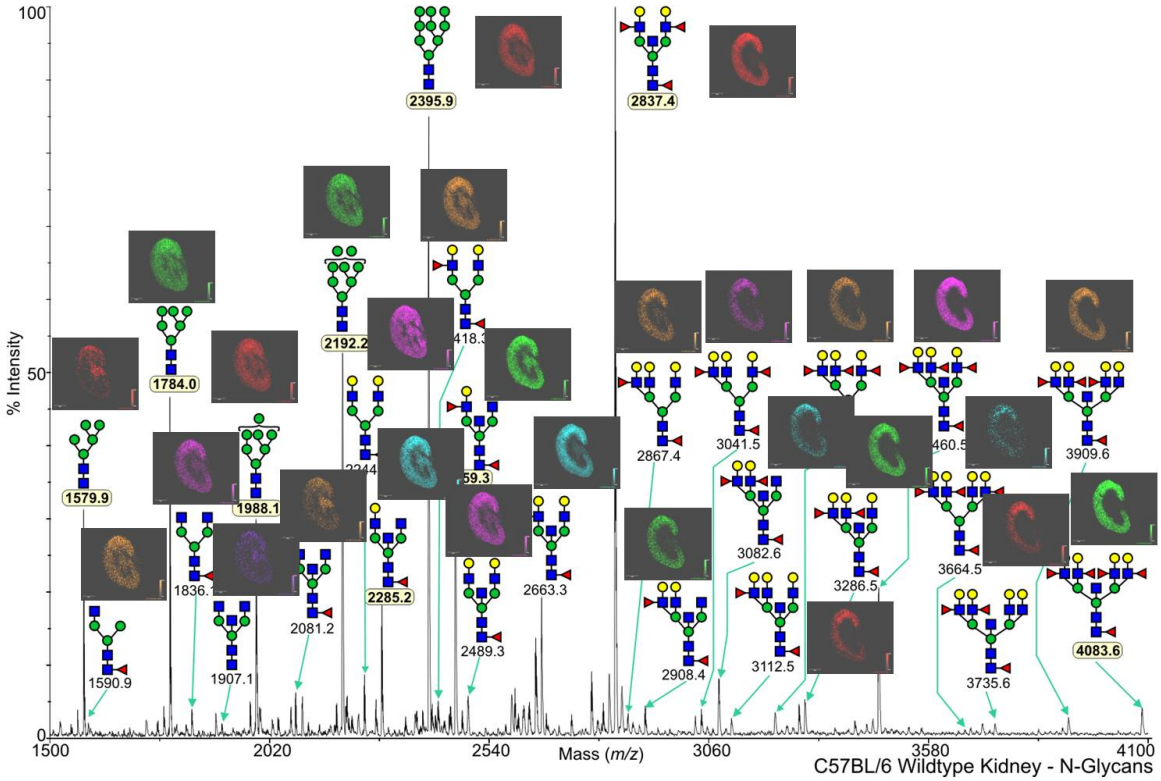


Figure 21. Permethylation of Extracted N-Glycans from a Mouse Kidney. Mouse kidney N-glycans were extracted from the imaging slide after PNGaseF application and digestion. Glycans were dried down and underwent permethylation as previously described. The permethylated m/z values were then compared to the permethylation data from the Consortium for Functional Glycomics mouse kidney database (www.functionalglycomics.org).



With Permission From CFG Core C Lead By Drs. Anne Dell and Stuart Haslem (www.functionalglycomics.org)

Figure 22. Correlation of Released N-Glycans with CFG Database.

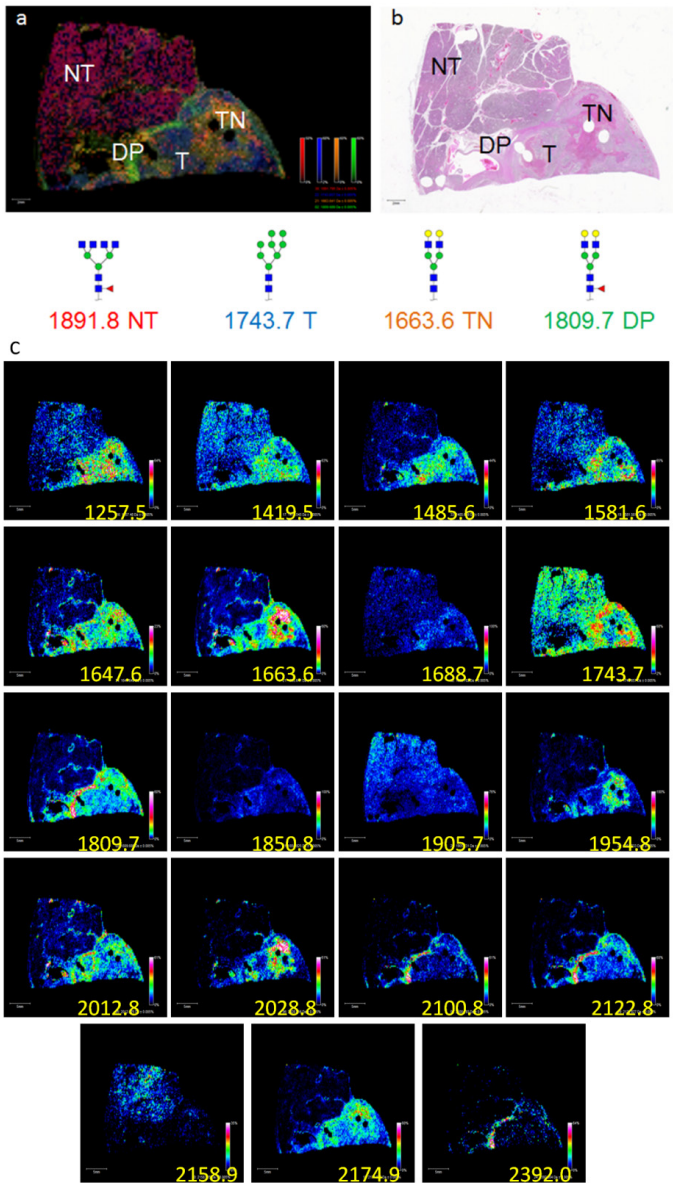


Figure 23. MALDI-IMS of a Human Pancreas FFPE Block. A FFPE block of pancreatic tissue from a human patient was cut at 5 µm prior to and selected for MALDI-IMS. Histopathology found four unique regions in the H&E of this tissue block. The tissue block contained tumor tissue, non-tumor tissue, fibroconnective tissue representing desmoplasia surrounding the tumor tissue, and necrotic tissue (b). MALDI-IMS was able to distinguish these four regions based off of specific ions after MALDI-IMS. M/z=1891.80 (red) is found in the non-tumor (NT) region of the pancreas and corresponds to Hex3dHex1HexNAc6, while m/z=1743.64 (blue) represents Hex8HexNAc2 and is predominant in the tumor region (T) of the tissue. Desmoplasia (DP) is represented by m/z=1809.69 (green) corresponding to Hex5dHex1HexNAc4. In the region where necrosis was identified (TN), m/z=1663.64 (orange) was elevated corresponding to Hex5HexNAc4. Image spectra were acquired at 200 µm raster. (c). Representative individual glycan images for the pancreatic FFPE tissue slice.

tissue (Figure 23b). A glycan of m/z = 1891.80 (red)/Hex3dHex1HexNAc6 was detected primarily in the non-tumor region of the pancreas, while a glycan of m/z = 1743.64 (blue)/Hex8HexNAc2 was predominant in the tumor region of the tissue. A region of desmoplasia surrounding the tumor region, an area of increased extracellular matrix proteins and myofibroblast-like cells resulting in a dense fibrous connective tissue, is represented by a glycan of m/z = 1809.69 (green)/Hex5dHex1HexNAc4 (433). A region of tumor necrosis is represented by a different glycan of m/z = 1663.64 (orange)/Hex5HexNAc4. Additional examples of tissue distributions of other individual glycan species are shown in Figure 23c.

A human prostate tissue block containing both tumor and non-tumor gland regions was also analyzed by MALDI-IMS. A heterogeneous N-glycan distribution reflective of the tissue histology was observed, and as an example of stroma and gland distributions, two glycan ions and two sub-regions within the tissue are highlighted in Figure 24. Distribution of glycans of m/z = 1663.56 (Hex5HexNAc4) and m/z = 1850.65 (Hex4dHex1HexNAc5) are shown in Figure 24b-c. A higher resolution tissue imaging analysis was done for selected regions as marked in the panel, with the H&E images (Figure 24d-f) highlighting stroma and gland substructures. In both instances, m/z = 1850.65 is present in both the stroma and glands, while m/z = 1663.56 is predominantly located in the stroma. An overlay of these two ions depicts the stroma as an orange color, demonstrating the presence of both red and green, while the glands are predominantly green. The distribution of other representative individual glycan ions is provided in Figure 25, including the distribution of high-mannose glycan species (Man5-Man9) associated with the heterogeneous tumor region in this tissue.

3.3.4.2 On-Tissue Glycan Fragmentation and Structural Composition

The glycan structures identified by imaging of the FFPE tissue blocks were assigned based on the comparison to permethylated species, glycan reference databases and previous studies (144). An on-tissue approach to further verify N-glycan structures was done using

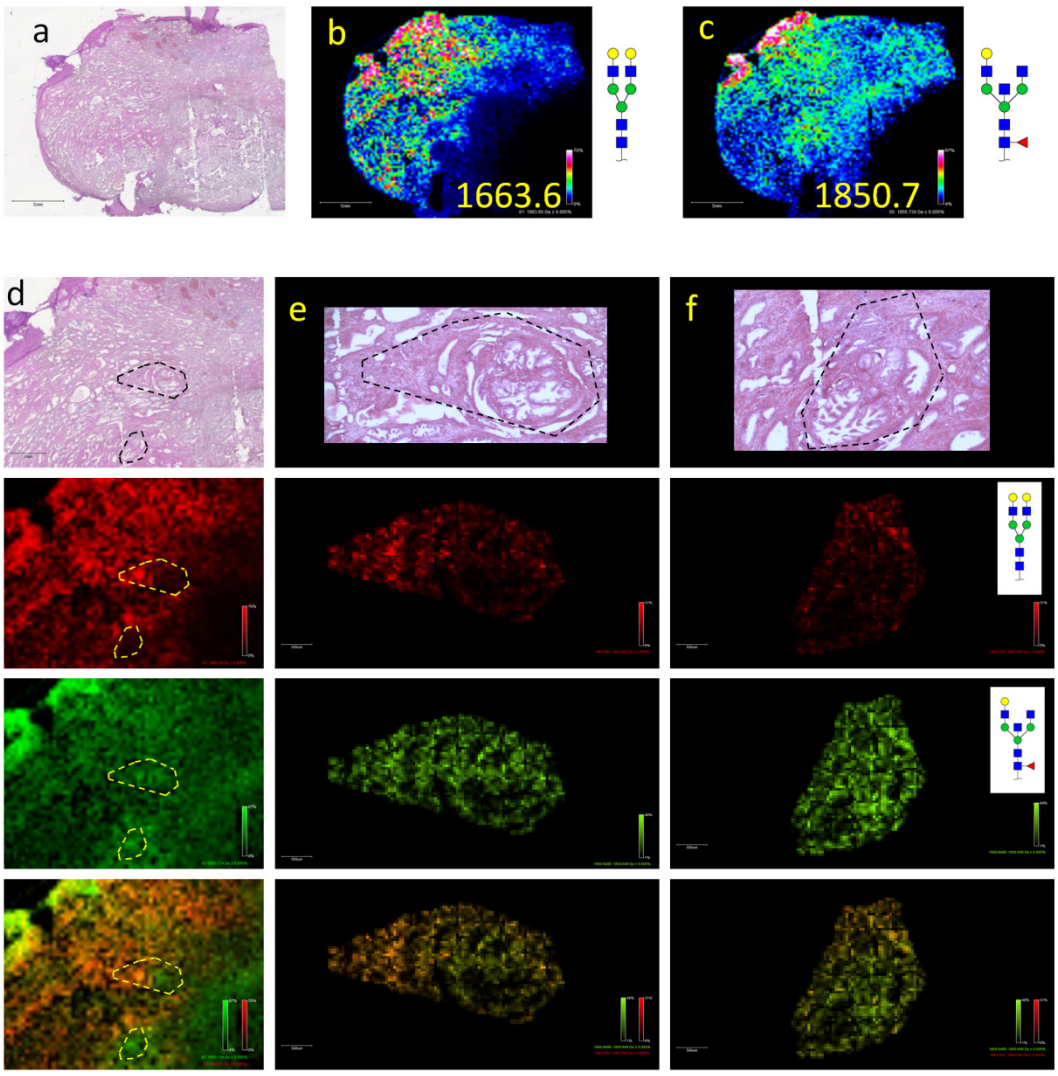


Figure 24. MALDI-IMS of a Human Prostate FFPE Tissue Block. An archived FFPE block of prostate tissue from a human patient was cut at 5 µm and prepared for MALDI-IMS glycan analysis, (a). H&E image. A global glycan imaging experiment performed with a raster of 225 µm demonstrated a heterogeneous expression of two glycan ions (b). at m/z=1663.56 and (c). m/z=1850.65. Stromal versus gland distribution were further assessed in a high resolution experiment at 50 µm raster (d–f). Column (d) indicates a 2× amplification of the H&E, and distribution of the same two glycans are shown at this magnification for m/z=1663.56 (red) and m/z=1850.65 (green), and an overlay image. Column (e) (enlargement of upper region shown in d). and (f) (enlargement of lower region shown in d), show two highlighted regions of stroma and glands enhanced at 10× resolution, with the same colors and glycans shown for column (d).

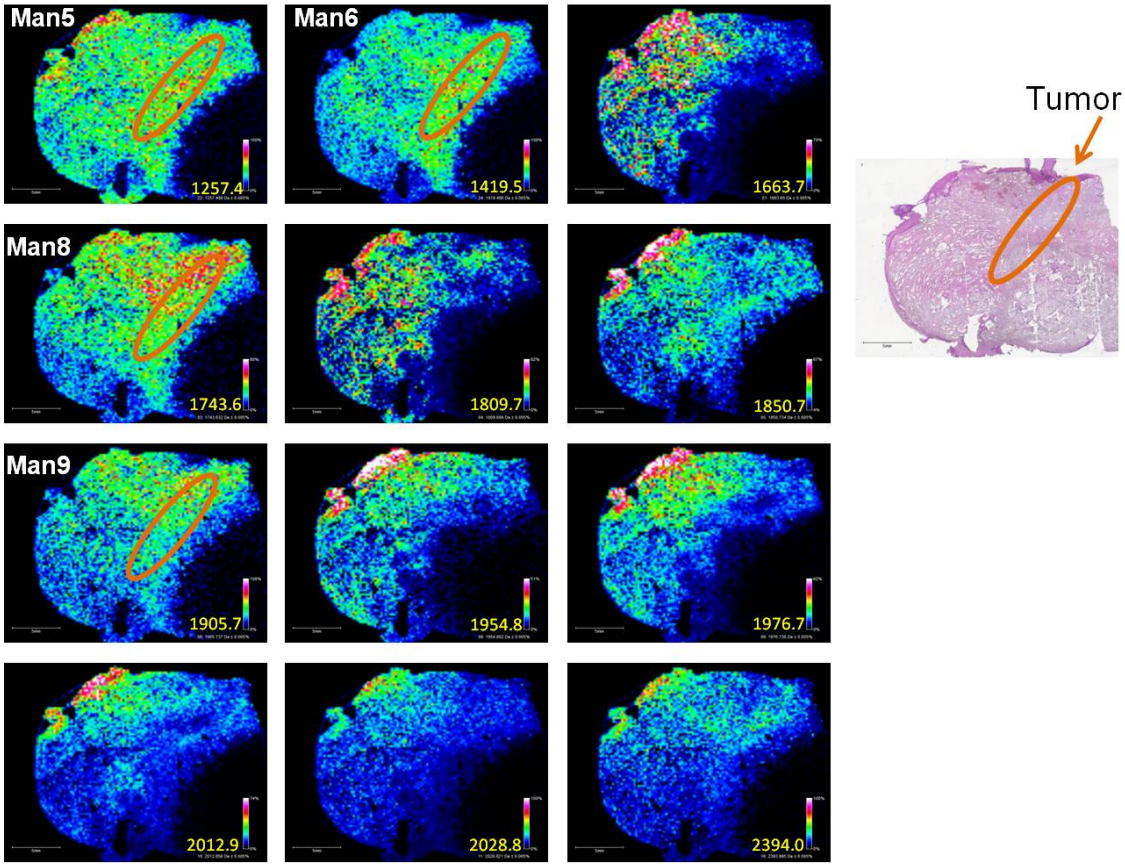


Figure 25. N-Glycan Panel in Prostate Tissue. A representative image of 12 glycoforms in a prostate tissue block, where high mannose glycoforms exhibited elevated abundance in the tumor tissue.

collision-induced dissociation (CID) directly on the human pancreatic tissue. Released native glycans from pancreatic cancer FFPE tissues were used as a source for on-tissue CID analysis, and a representative MALDI spectra of these glycans is shown in Figure 26a. For comparison, a Hex5HexNAc4 (m/z = 1663.6) purified standard (also termed NA2) was spotted on a stainless steel MALDI target plate and fragmented by CID, generating a robust fragmentation pattern of glycans for this ion as previously reported by Harvey et al (247). The same glycan ion was abundant in pancreatic tissue after PNGaseF release of N-glycans (Figure 27) and was selected for CID. As shown in Figure 5b, the CID fragmentation pattern of m/z 1663.6 in pancreatic tissue was the same as the N-glycan standard, confirming detection of NA2 directly in pancreatic tissue (Figure 27). Mass shifts due to loss of individual sugar ions were detected, such as Hex (resulting in m/z = 1502.5), HexNAc (resulting in m/z 1460.5), and Hex + HexNAc (resulting in m/z = 1298.5) (Figure 26). An ion at m/z = 712.2, which has been previously characterized as the sodium adduct of Hex3HexNAc1, was also detected (247). The structures of 13 other glycan ions were confirmed using this CID approach, and additional fragmentation data and spectra are provided in Figure 26.

3.3.4.3 Glycan MALDI-IMS of a Hepatocellular Carcinoma Tissue Microarray

The ability to perform N-glycan analysis on FFPE tissues potentially enables the analysis of multiple FFPE tissue cores in a TMA format. Initial experiments were performed using a commercially available hepatocellular carcinoma (HCC) TMA (BioChain) consisting of samples from 16 individual patients, with two tumor tissue cores and one non-tumor tissue core per patient (Figure 28). Additional patient data are provided in Sup Sheet 4. Glycan MALDI-IMS was done as described for the other FFPE tissues, and imaging data for two representative glycan ions at m/z = 2393.92 (Hex7HexNAc6) and m/z = 1743.62 (Hex8HexNAc2) are shown in Figure 28. Analysis of the cumulative MALDI spectra and detected ions for each tissue core were processed and compared using an in-house bioinformatic workflow followed by statistical

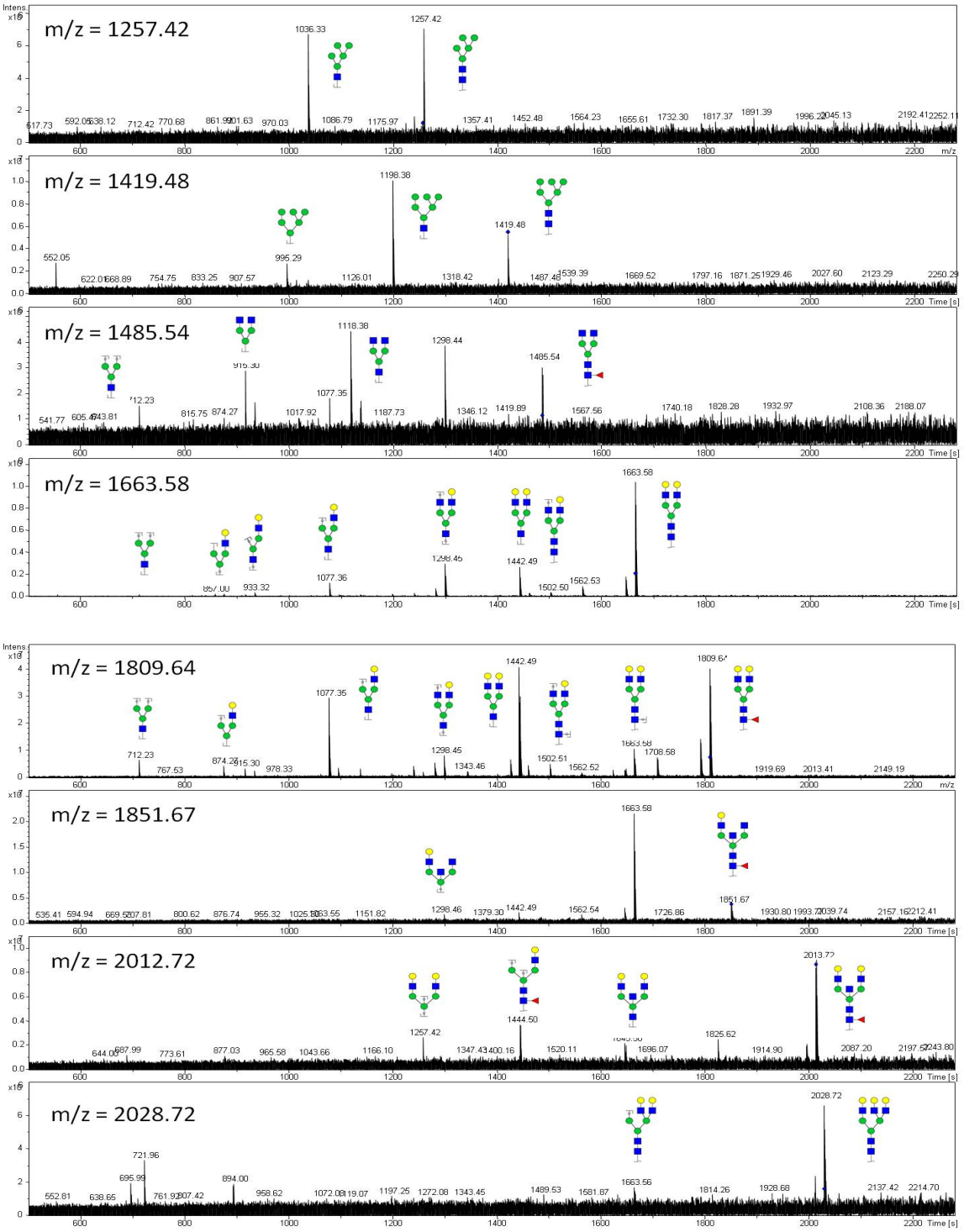


Figure 26. CID of N-Glycans from Human Pancreas Tissue. Representative CID spectra from 8 glycans selected for on tissue fragmentation are provided in an annotated format.

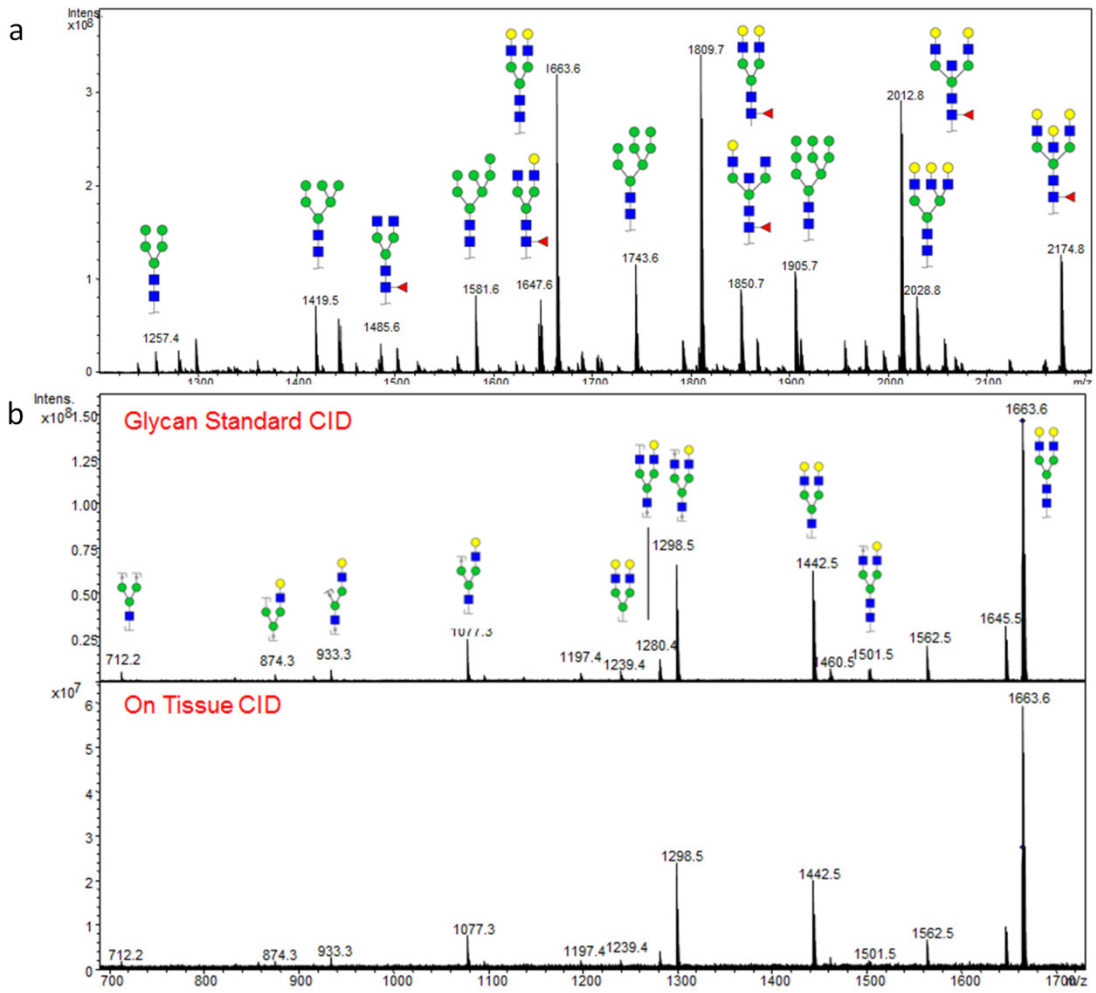


Figure 27. Comparison of the Fragmentation Pattern of a Glycan Standard with the same Ion on Tissue. (a). A representative MALDI spectra for native N-linked glycans from pancreatic cancer FFPE tissue. (b). NA2 glycan standard (m/z=1663.6) was fragmented using CID, revealing a variety of cleavages across glycosidic bonds as demonstrated in the spectrum (a). When the same ion was fragmented on the pancreatic tissue, the fragmentation pattern was the same, verifying that we were detecting Hex5HexNAc4 in the human pancreas.

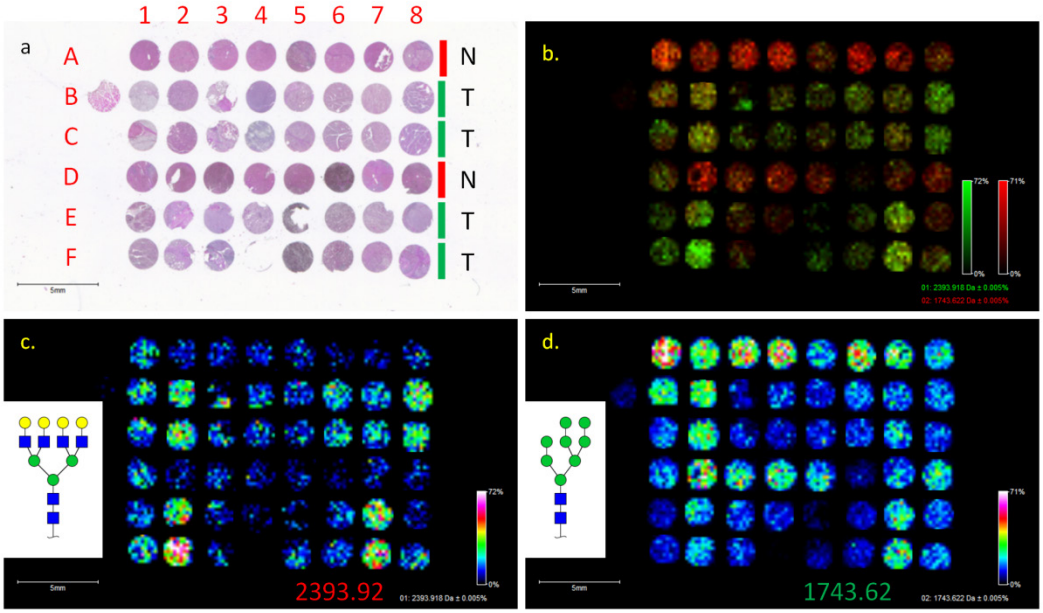


Figure 28. N-Glycan Imaging of a Liver TMA. A liver TMA purchased by BioChain consisting of 2 tumor tissue cores and one normal tissue core from 16 patients was imaged (200 µm raster). The H&E (a) provides the TMA location (red letters and numbers) and classifies whether the row is tumor (green bar) or non-tumor (red bar). M/z=2393.95 (c) and m/z 1743.64 (d) were able to distinguish between hepatocellular carcinoma and uninvolved liver tissue. An overlay of these ion demonstrates that m/z=2393.95 is elevated in tumor tissue and m/z=1743.64 is elevated in normal tissue (b).

analysis. Of the 176 identified ions from the HCC TMA, 132 were increased in tumor cores, and 83 ions had a p-value <0.05. Interestingly 78 (94%) of the significantly different ions were elevated in tumor cores. After cross-referencing this list of 176 ions with glycans presented in this paper and our previous study, 26 N-glycans of high-confidence structure determinations were selected, listed in Table 2 (431). Of these 26 known glycans, ion intensities of 13 species were significantly different in tumor and normal tissue ($p<0.05$), and 21 were increased in tumor relative to normal. FlexImaging was then used to demonstrate the distribution and relative ion intensities of each glycan across the TMA (Figure 29). Additionally, ROC curves were used to evaluate how well each of the glycan ion intensities discriminates tumor versus non-tumor. Of the 176 identified ions, 61 had area under the ROC curve (AuROC) >0.80, indicating they are strong classifiers. For two glycans at m/z = 2393.95 (Hex7HexNAc6) and m/z = 1743.64 (Hex8HexNAc2), both had an AuROC > 0.80 and a p-value < 0.05, with m/z 2393.95 being elevated in tumor tissue and m/z 1743.64 being elevated in non-tumor tissue, as demonstrated by the log2-fold change value (tumor/non-tumor) (Figure 28). In the overlay (Figure 28b), tumor tissue is predominantly green and non-tumor tissue is predominantly red, confirming results from our statistical analysis. This data, although from limited numbers of samples, demonstrates the potential ability of a panel of glycans to be used to accurately discriminate cell types or outcomes on a TMA by MALDI-IMS.

3.3.5 Discussion

Multiple N-linked glycans can be directly profiled from FFPE tissue blocks and TMAs while maintaining intact architecture. The basic methodology, which mirrors that of MALDI-IMS analysis of peptides in FFPE tissues and TMAs, requires deparaffinization and antigen retrieval prior to PNGaseF application (144,429,430). The ability to adapt the N-glycan imaging method originally designed for fresh/frozen tissues to encompass FFPE tissue and TMA blocks increases the scope and speed of glycan-based studies that can be performed in tissues (431). In initial

Table 2. Comparison of Tumor and Non-Tumor Glycans Detected in Hepatocellular Carcinoma Tissue Microarrays.

m/z	AuROC	log2-fold change	p-val
1866.76	0.895	3.045	<.001
2393.95	0.879	1.819	<.001
2378.01	0.869	3.495	<.001
1743.64	0.831	-0.967	<.001
1850.78	0.827	1.953	<.001
1257.42	0.821	0.78	<.001
1501.6	0.813	2.223	0.001
2686.02	0.764	4.316	0.003
1905.71	0.742	-0.572	0.007
1298.44	0.742	0.91	0.007
2012.82	0.734	1.165	0.01
2320.89	0.707	2.146	0.022
2540.03	0.707	1.493	0.022
3271.15	0.657	2.004	0.082
2028.74	0.657	0.36	0.082
1647.62	0.649	0.683	0.099
1581.57	0.635	-0.19	0.135
2174.89	0.633	0.598	0.141
1976.71	0.617	0.398	0.197
2100.77	0.597	0.295	0.266
1954.79	0.593	-0.114	0.307
1419.48	0.577	-0.076	0.4
1809.67	0.548	0.973	0.598
1485.54	0.488	0.149	0.902
1663.66	0.49	0.255	0.92
1282.46	0.494	0.703	0.991

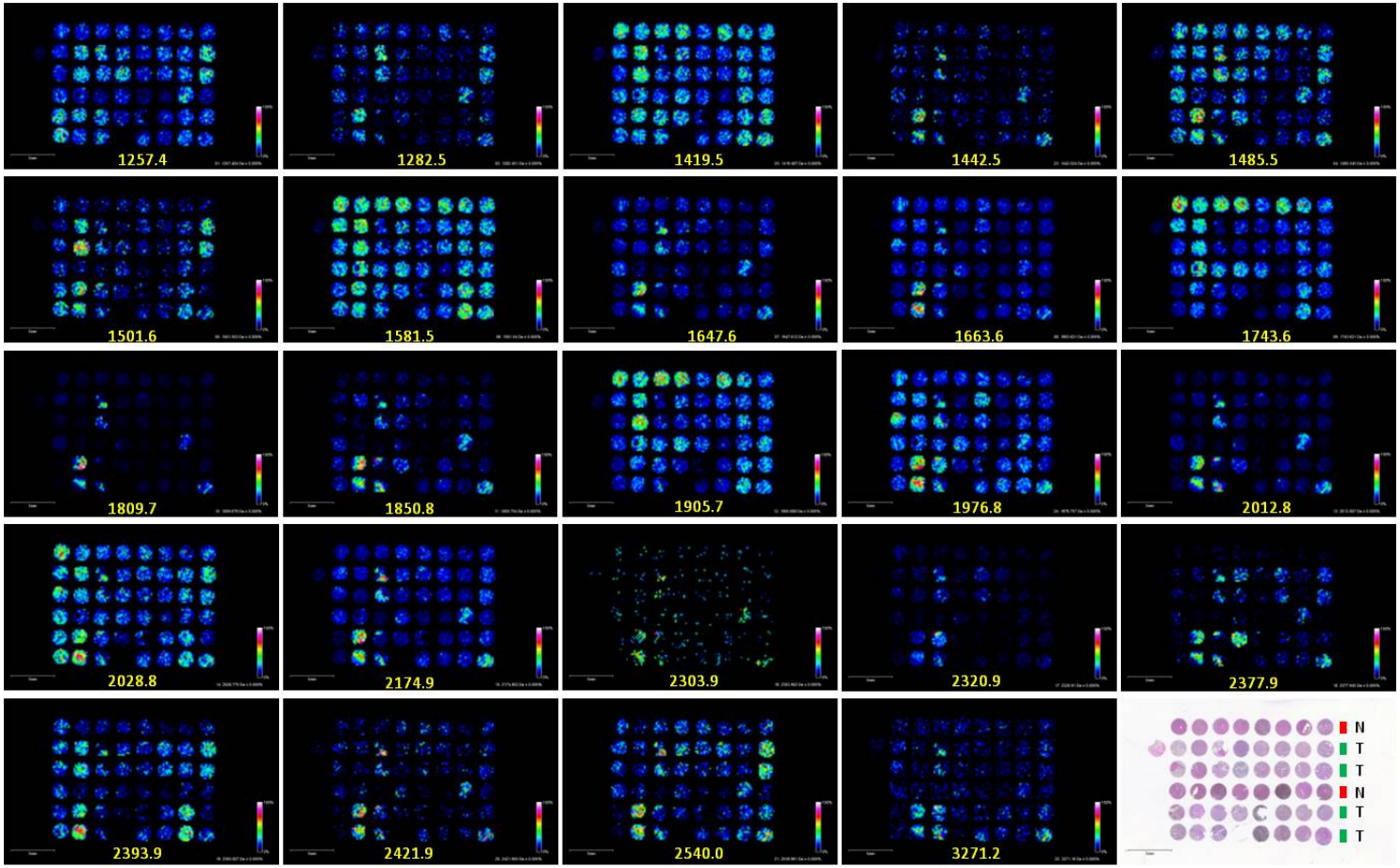


Figure 29. Panel of N-Glycans Detected in the HCC TMA.

studies of formalin-fixed mouse kidney slices, the MALDI-IMS workflow successfully identified all 28 of the glycans in the mouse kidney database provided by the Consortium for Functional Glycomics. Many of the structures of these glycans were verified by permethylation (Figure 21) and CID experiments (Figure 26). As observed with the mouse brain, these glycans were not homogenously present across the entire mouse kidney slice, but were either predominantly located in the cortex, or distributed across the cortex and medulla (Figure 20) (431). This unique distribution of N-glycans associated with tissue sub-structure or disease status was also observed in human pancreas and prostate tissue slices. In the pancreas, an overlay of four different glycans was able to map the normal pancreas tissue, tumor pancreas tissue, a region of desmoplasia, and a necrotic region (Figure 23a). Similarly, an overlay of two glycans could distinguish between prostate stroma and glands (Figure 24).

In general, the peak intensities of PNGaseF-released glycans in the FFPE tissues seems to be more intense than that obtained with fresh/frozen tissue sections. This may be a result of the more extensive heating and washing steps required in the deparaffinization and rehydration steps. It is this increased detection sensitivity that facilitated CID fragmentation of N-glycans directly from the tissue (Figure 26, Figure 27b). Under the conditions used, CID generated mainly fragments across the glycosidic bonds, which were useful in characterizing that the structure was an intact hexose or HexNAc. This did not provide any information regarding anomeric linkages between sugar residues. The amount of fragmentation observed was directly related to the relative intensity of each parent N-glycan ion, and inversely related to the mass of the parent ion observed. This is typified by the extensive fragmentation of two glycans of m/z = 1663.50 and m/z = 1809.64 Figure 26, Figure 27b).

One drawback to using FFPE tissues is residual polymer from the paraffin block adjacent to the tissue. Detection of this polymer is more predominant in the lower mass range of the imaging runs, and can overlap with potential glycan masses, complicating detection and further

statistical analysis. This polymer can be observed in the average spectra of the mouse kidney tissue after PNGase application (Figure 19a) from m/z = 1250-1300, 1450-1500, and 1650-1700. An additional key to distinguishing polymer peaks is the analysis of spectra from the non-PNGaseF treated control tissues. Particularly for the TMA format, the ion selection program that we report can detect and account for polymer peaks relative to glycan ions. These polymer peaks seem to vary in terms of intensity compared to N-glycan ions depending on what tissue is being used. It is possible that this variation is a function of different formalin formulations, variations is tissue processing (i.e. amount of time in formalin), storage time or variations in the tissue itself (419,420). These considerations will be further monitored and evaluated as more glycans from FFPE tissues are analyzed.

In relation to potential cancer diagnostic applications, the most significant aspect to developing a method to image N-glycans on FFPE tissue blocks could be the ability to use TMAs for high-throughput glycan-based experiments. Not only does the method increase the number of tumor samples that can be analyzed in one experiment, but it could also be used to compare the glycans detected in a TMA core versus the larger source FFPE tissue. N-glycan MALDI-IMS of the HCC TMA (Figure 28) is provided as an example, but we have already obtained initial glycan profiling data from TMAs representing prostate, kidney, lung, breast, colon and pancreatic cancers. In the HCC TMA, a statistically significant increase in tetra-antennary N-glycan (m/z = 2393.95) and decrease in Man-8 glycan (m/z = 1743.64) was detected in HCC cores compared to adjacent non-tumor tissue (Figure 28 and Table 2). The tetraantennary N-glycan has been previously demonstrated to be elevated in HCC compared to matched adjacent non-tumor tissue by Mehta et al. (434). Continued investigations will be performed on whether these two ions can distinguish between matched HCC and non-tumor tissues in other HCC TMAs. Our data analysis identified a total of 176 ions in the tissue, with the majority of significantly different ions being increased in HCC relative to non-tumor tissue, including 21 known or previously identified

glycans. It is unclear how this trend of increased glycan levels relates specifically to tumor related biochemical changes, though the role of glycosylation in tumor development is well documented. Future work will also focus on determining the identity of the remaining ions to distinguish other glycan species from the aforementioned polymer peak contaminants.

Currently, MALDI-IMS provides a new approach to effectively visualize and evaluate N-glycan localization in tissue sections. It does not solve the known limitations of MALDI analysis of underivatized glycans like loss of sialic acids, nor does it provide anomeric linkage information for N-glycan structure. Established tandem mass spectrometry methods of glycan extraction, modification and fragmentation are more capable of providing this structural information. Combining the glycan tissue maps generated by MALDI-IMS to target regions of interest for further tandem mass spectrometry analysis of glycans could be a new synergistic approach to more effectively identify tumor-associated glycans and glycoproteins in situ. Use of other glycosidases like sialidase, as we have previously reported, or fucosidases, could further extend the utility of the combined methods (431). Overall, the ability to effectively profile N-glycans on FFPE tissue blocks and TMAs provides new opportunities to evaluate glycan profiles associated with disease status.

3.4 Disease-Related Applications of and Limitations of MALDI-IMS

3.4.1 Introduction

Hepatocellular carcinoma (HCC), a subtype of liver cancer that accounts for approximately two-thirds of all liver cancers, is among the most common and aggressive malignancies worldwide (435). As incidence rates continue to rise, early detection disease biomarkers are highly sought after and offer the potential to drastically improve patient outcome(436–439). Both proteomic and glycomic approaches have led to significant advances in HCC biomarker research. Proteomic analysis has identified several tumor pathways that are altered in HCC (440–443). Additionally, glycoproteins such as peroxiredoxin 3, osteopontin, and alpha fetoprotein (AFP) have been identified as potential HCC biomarkers (444–448). However, these proteins display limited sensitivity in the background of liver cirrhosis or benign diseases. More recently, it was observed that α1,6 core fucosylated AFP offers improved specificity for HCC than AFP alone (350,449,450). Thus, core fucosylated AFP, assessed through the AFP-L3 test, is the only test approved by the United States Food and Drug Administration for the detection of hepatocellular carcinoma. Elevated levels of core fucosylation are not specific for AFP, but are also observed on transferrin and alpha-1-antitrypsin among other proteins (451–455). Total serum glycomic studies have also identified increases in core α1,6-linked and α1,3-linked outer-arm fucosylation, N-glycan branching, and sialylation in HCC samples (219,434,456,457).

As nearly all current cancer biomarkers are glycoproteins or carbohydrate antigens, and as highlighted for HCC, a global analysis of N-glycans from HCC tissue sections using MALDI-imaging mass spectrometry (IMS) could extend and complement the continued development of N-glycan biomarkers associated with HCC (416,458). We have recently developed a MALDI-IMS approach for profiling N-glycans directly in frozen and formalin-fixed paraffin-embedded

(FFPE) tissues (431,459). The method relies on the spraying of a molecular coating of PNGaseF and on-tissue cleavage of N-glycans prior to matrix application. Relative to other classes of biomolecules targeted by MALDI-IMS, the method offers ready identification of the individual glycan species and creation of structural reference databases (459). In this report, the utility of the method is demonstrated for profiling multiple N-glycans using different FFPE tissue slices of HCC. Current data analysis approaches are also summarized, as well as use of a novel derivatization method to stabilize sialic acids in a linkage-dependent/differentiating manner, and therefore better characterize larger, sialic acid containing N-glycans (237). As the methodology is still evolving, areas to improve glycan detection by MALDI-IMS and structural characterization are discussed.

3.4.2 Materials and Methods

3.4.2.1 Materials

Trifluoroacetic acid, α-cyano-4-hydroxycinnamic acid (CHCA), sodium hydroxide (NaOH), 1-hydroxybenzotriazole hydrate (HOBt), and trypsin were obtained from Sigma-Aldrich (St. Louis, MO). Ammonium bicarbonate, HPLC grade methanol, ethanol, acetonitrile, xylene and water were obtained from Fisher Scientific (Pittsburgh, PA). 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) was obtained from Oakwood Chemical (West Columbia, SC). ITO slides were purchased from Bruker Daltonics (Billerica, MA) and Tissue Tack microscope slides were purchased from Polysciences, Inc (Warrington, PA). Citraconic anhydride for antigen retrieval was from Thermo Scientific (Bellefonte, PA). Recombinant Peptide N-Glycosidase F (PNGaseF) from *Flavobacterium meningosepticum* was expressed and purified as previously described (431), and is available commercially as PNGase F Prime™ from Bulldog Bio (Portsmouth, NH).

3.4.2.2 FFPE Tissues

All but one tissue sections used in this study were de-identified and obtained commercially from BioChain. These include a slide with two FFPE tissue sections from a 60 year old female with a poorly differentiated hepatocellular carcinoma; patient matched tumor and normal FFPE tissue sections from a 53 year old male with hepatocellular carcinoma, and a patient matched renal cell carcinoma tumor and lymph node tissue with renal cell carcinoma metastasis from a 28 year old female. A subset of a commercially available renal tissue microarray (TMA) from BioChain was also analyzed, where two cores from each patient (oriented vertically) were present. One de-identified pancreas tissue was obtained by MUSC and was determined to be not human research classifications by the respective Institutional Review Boards at MUSC.

3.4.2.3 Washes for Deparaffinization and Rehydration

Slide preparation proceeded as described in our previous paper (459). Briefly, FFPE tissue sections not acquired precut from BioChain, were sectioned at 5µm and mounted on slides compatible with the Bruker slide adaptor. All slides were heated at 60°C for 1hr to ensure tissue adhesion to the slide. After cooling, the slide was deparaffinized by washing with xylene and rehydrated in a series of ethanol and water washes. Citraconic anhydride (Thermo) was used as the antigen retrieval buffer and the retrieval process took place over 25 minutes in a vegetable steamer. After allowing the buffer to cool, the buffer was gradually exchanged to 100% water. The slide was then desiccated to dryness prior to enzymatic digestion.

3.4.2.4 MALDI Imaging Workflow

An ImagePrep spray station (Bruker Daltonics) was used to coat the slide with a 0.2mL aqueous solution of PNGaseF (20µg total/slide) as previously described (431). As a negative control, adjacent control tissue slices were shielded from PNGaseF application by covering the tissue section with a glass slide. Digestion occurred in a humidified chamber at 37°C for 2hr.

Slides were desiccated prior to α-cyano-4-hydroxycinnamic acid matrix application (0.021g CHCA in 3mL 50% acetonitrile/50% water and 12µL 25% TFA) using the ImagePrep sprayer. Released glycan ions were detected using a SolariX dual source 7T FTICR mass spectrometer (Bruker Daltonics) (*m/z* 690-5000) with a SmartBeam II laser operating at 1000 Hz, a laser spot size of 25µm. Following MS analysis, data was loaded into FlexImaging Software focusing on the range m/z = 1000-4000 and reduced to 0.95 ICR Reduction Noise Threshold. FlexImaging 4.0 (Bruker Daltonics) was used to generate images of differentially expressed glycans. Observed glycans were searched against the glycan database generated using GlycoWorkbench (432). Presented glycan structures were generated in GlycoWorkbench and represent putative structures determined by combinations of accurate m/z and off-slide derivatization experiments.

3.4.2.5 Ethyl Esterification

N-glycans were extracted from slides as described previously and dried by vacuum centrifugation (431,459). The ethyl esterification protocol, including the modification and enrichment, was adapted from Reiding et al. (237). Briefly, 2µL water and 40µL 0.25M HOBt/EDC were added to dried glycans followed by incubation at 37°C for 1hr. 40µL acetonitrile was added and the mixture was placed at -20°C for 20 minutes. Glycans were enriched using cotton-HILIC tips according to Selman et al. Shortly, cotton wool composed of 100% cotton (Assured, Rio Ranch, NM) was inserted in 20µL tips and equilibrated with 10µL water three times followed by 10µL 85% ACN three times. Samples were loaded and unbound material removed by washing three times with 10µL 85% acetonitrile with and without 1% TFA, respectively. Tip-bound glycans were eluted in 10µL water. Enriched and modified glycans were spotted on an Anchorchip MALDI plate (Bruker Daltonics) with 2,5-dihydroxybenzoic acid (DHB) at a concentration of 5mg/mL in 50% ACN/50% water/1mM NaOH. Ethanol was then used for recrystallization.

3.4.3 Results

3.4.3.1 Influence of Histopathology on MALDI-IMS of N-glycans

Initial MALDI-IMS of a commercially available FFPE HCC tissue of complex histopathology demonstrated the specific release of N-glycans following PNGaseF digestion and the ability of released glycans to distinguish tissue subtypes and pathologies. Two serial sections were prepared identically with the exception that one tissue received PNGaseF while the other was shielded from enzyme application to serve as a negative control. Average spectra of each tissue region were obtained from FlexImaging and revealed a robust signal increase following PNGaseF digestion (Figure 30a) compared to the control tissue (Figure 30b). Putative glycan structures were annotated using GlycoWorkbench based on mass accuracy and previous studies, but no information is available regarding specific anomeric linkages (431,459). All glycan structures reported herein are the $[M + Na]^+$ adducts unless otherwise noted. Peaks corresponding to three glycans, Hex5HexNAc4NeuAc1 (m/z = 2100.759, blue), Hex5HexNAc4 (m/z = 1663.582, red) and Hex9HexNAc2 (m/z = 1905.612, green) are displayed in an overlay image (Figure 30c). These glycans map to the tissue histopathology marked on the H&E stain (Figure 30d). Analysis of the H&E stain revealed three predominant tissue morphologies; necrosis (outlined in red), HCC tissue (outlined in green) and fibroconnective tissue (outlined in blue). Images of individual N-glycans demonstrate the propensity of glycans to be elevated in one tissue type (Figure 34). Interestingly, high-mannose glycans were elevated in the tumor tissue, while fucosylated or sialylated/fucosylated complex biantennary glycans were elevated in the fibroconnective tissue.

While the initial experiment demonstrated the ability of N-glycan localization to define pathology in HCC tissue blocks, the ability to distinguish tumor vs. normal tissue is of more clinical significance. Analysis of a patient matched tumor and normal FFPE tissue sample

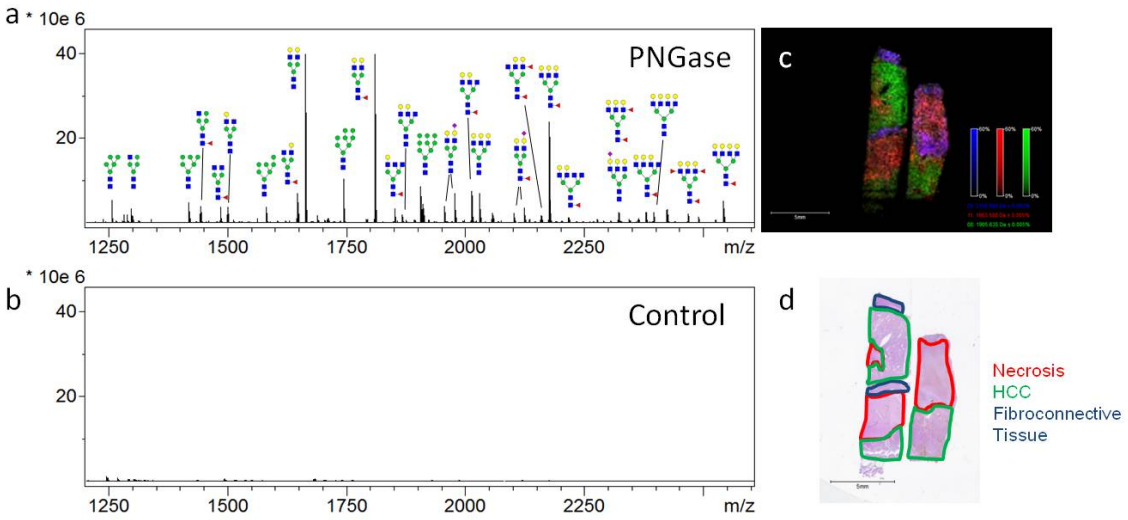


Figure 30. PNGaseF Releases N-linked Glycans from HCC Tissue in a Spatially Defined Manner. A HCC tissue slide (Biochain) containing two serial slices of HCC tissue was prepared from MALDI-IMS as described above. One section was shielded from enzyme application to serve as a negative control. Robust signal correlating to released N-linked glycans was observed in the average spectra of the tissue receiving PNGaseF application (a). Similar signatures were not observed in the control tissue (b). Three select ions were selected in FlexImaging and displayed as an overlay image. Hex5HexNAc4NeuAc (m/z = 2100.7, blue), Hex5HexNAc4 (m/z = 1663.6, red) and Hex9HexNAc2 (m/z = 1905.6, green) each display unique localization within the tissue slice (c). Upon histological analysis of an H&E stain (d), these regions were identified to be fibroconnective tissue, necrotic tissue, and HCC tissue respectively.

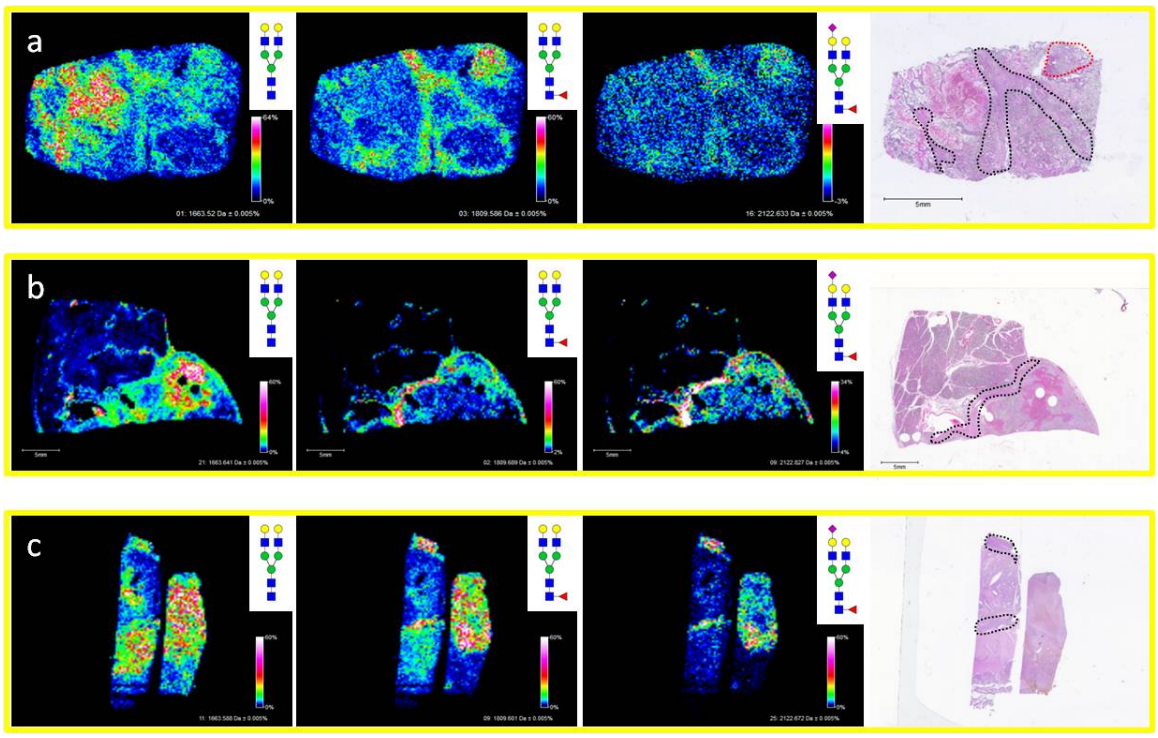


Figure 31. Diverse Organs with Similar Properties Display Conserved Glycan Patterns. Fibrous tissue from three diverse tissue types were imaged and displayed similar N-glycosylation properties. Images from lymph tissue with a ccRCC metastasis (a), pancreatic cancer tissue with complex histopathology (b), and liver tissue with complex histopathology (c) are provided. In each tissue, the ions corresponding to Hex5HexNAc4, Hex5dHex1HexNAc4, and Hex5dHex1HexNAc4NeuAc1 display similar localization patterns with each being elevated in the fibrous tissue regions. Hex5HexNAc4 and Hex5dHex1HexNAc4 are represented by the sodium adduct of the native glycan, while Hex5dHex1HexNAc4NeuAc1 is represented as a doubly sodiated adduct of the native glycan. Fibroconnective tissue regions are outlined in black on the H&Es while fibrotic ccRCC tumor tissue is outlined in red.

revealed overall glycan heterogeneity between the two tissues. While Hex8HexNAc2 (m/z = 1743.565, red) is present in both the normal and tumor sections, Hex7HexNAc6 (m/z = 2393.854, green) is largely absent in the normal tissue (Figure 32a,b). This observation is evident in the image overlay, where the normal tissue image is red in color due to the presence of Hex8HexNAc2 and the absence of Hex7HexNAc6, while the tumor tissue is yellow due to the presence of both Hex7HexNAc6 and Hex8HexNAc2 (Figure 32c). This finding is consistent with our previous studies made in the analysis of an HCC TMA (459). Interestingly, while Hex8HexNAc2 is present in both the matched tumor and normal tissues and is elevated in the normal HCC tissue, it is elevated in the tumor tissue compared to necrotic and fibroconnective tissue regions (Figure 30, Figure 34). This trend emphasizes the importance of histological analysis when utilizing MALDI-IMS as a technique.

3.4.3.2 Similarities of Glycan Distribution Across Tissue Types.

The distribution of N-glycans in different fibroconnective tissues from other sources besides HCC tissues was further assessed. Lymph tissue with a clear cell renal cell carcinoma (ccRCC) metastasis (Figure 31a), pancreatic cancer and adjacent normal tissue (Figure 31b), and the hepatocellular carcinoma tissue presented earlier (Figure 31c), all with regions of fibrous and/or fibroconnective tissue, were analyzed by MALDI-IMS for N-glycans. The regions of fibroconnective tissue (outlined in black) and fibrous tumor tissue (outlined in red) are defined on the H&E image for each tissue (Figure 31). In each tissue, three biantennary N-glycans were commonly detected at higher levels in the fibrous/ fibroconnective tissue compared to adjacent regions: Hex5HexNAc4 (m/z = 1663.582), Hex5dHex1HexNAc4 (m/z = 1809.661), and Hex5dHex1HexNAc4NeuAc1 (m/z = 2122.720; $[M - H + 2Na]^+$). In all tissues, the Hex5dHex1HexNAc4NeuAc1 glycan had the highest specificity for fibrous tissue regions. The Hex5HexNAc4 was detected at greater signal intensity in more regions of the tissues, and we

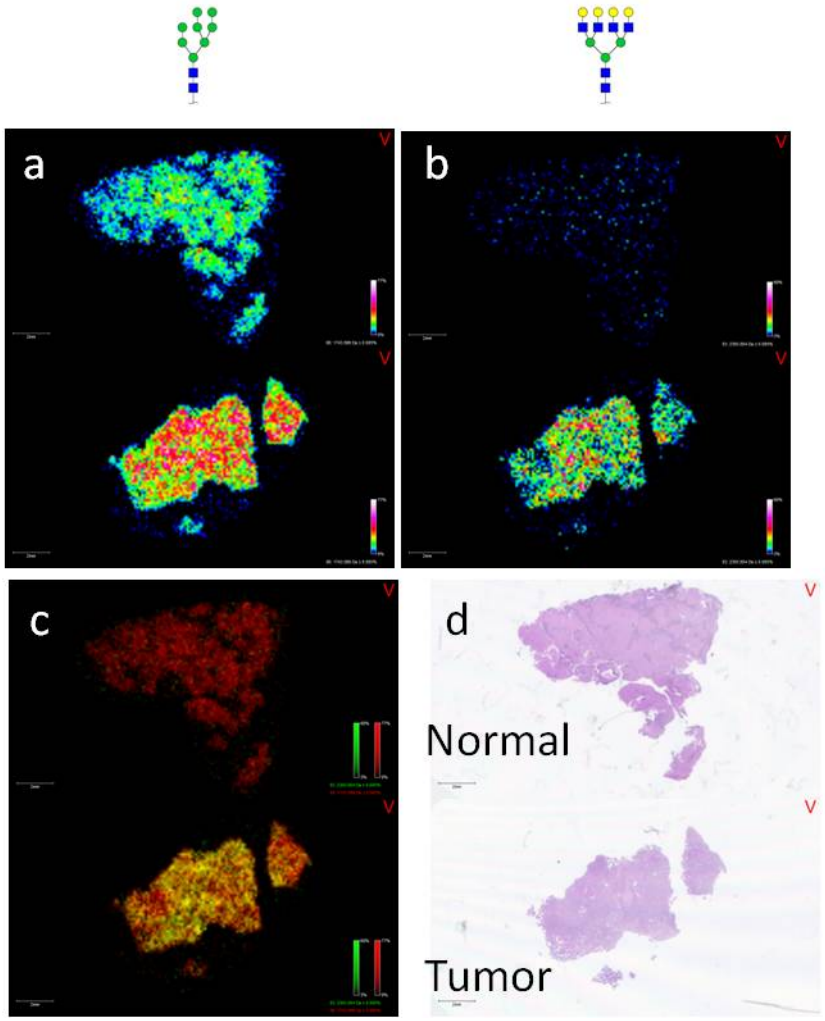


Figure 32. MALDI-IMS Reveals N-glycans that can Distinguish Tumor from Normal in Matched HCC Tissue Samples. Matched HCC tumor and normal tissue (biochain) were analyzed by MALDI-IMS after PNGaseF release of N-glycans. M/z = 1743.56 (Hex8HexNAc2) was present in both tissues (a), while m/z = 2393.85 (Hex7HexNAc6) was primarily in the tumor tissue only (b). An overview of these two glycan species where m/z = 1743.56 is displayed in red and m/z = 2393.85 is displayed in green confirms this finding. In the overlay, the normal tissue is primarily red but the tumor tissue is yellow in color as a result of both glycan species being present (c). H&E staining of both tissues is provided (d).

have also noticed that this glycan is likely a marker of tissue regions where blood is present, illustrated in Figure 31b.

3.4.3.3 Validation of N-Glycan Structures by Off-Tissue Glycomic Analysis

MALDI ionization of glycans often results in the loss of labile sialic acids from glycans. We performed a recently introduced glycan modification approach, termed ethyl esterification, to assess whether the loss of sialic acids is occurring in MALDI-IMS experiments. This procedure results in lactonization of α2,3-linked sialic acids and ethyl esterification of α2,6-linked sialic acids, resulting in a 46.04 Da mass shift (237). Following ethyl esterification of glycans, glycans are enriched using cotton-HILIC tips. Glycans from a HCC tissue were retrieved from the tissue, subjected to the ethyl esterification protocol, and spotted on an anchorchip MALDI plate. An annotated spectrum is provided (Figure 33). A list of 33 representative N-glycans present in the MALDI-IMS and/or ethyl esterification experiments, as well as the glycans theoretical m/z value is provided in Table 3. The results demonstrate a close overlap of observed m/z values for both approaches with theoretical m/z values obtained from GlycoWorkbench. Additionally, a majority of N-glycans were present in both the MALDI-IMS and off-tissue ethyl esterification experiments. However, when comparing the ethyl esterification average spectrum (Figure 33) with the MALDI-IMS average spectrum (Figure 30a), it is evident that complex and sialylated N-glycans are more predominant after ethyl esterification than in MALDI-IMS: an example being Hex5HexNAc4NeuAc1 which has the highest signal intensity in the ethyl esterification spectrum (m/z = 1936.692 & m/z = 1982.733), but is only detected as minor peaks in the imaging average spectrum (m/z = 1954.589 & m/z = 1976.707). This result may be attributed to a reduction in signal intensity of sialylated N-glycans in MALDI-IMS due to the loss of sialic acid upon in-source decay, reflected in the comparatively higher level of detection of Hex5HexNAc4 (m/z = 1663.582) in the MALDI-IMS average spectrum.

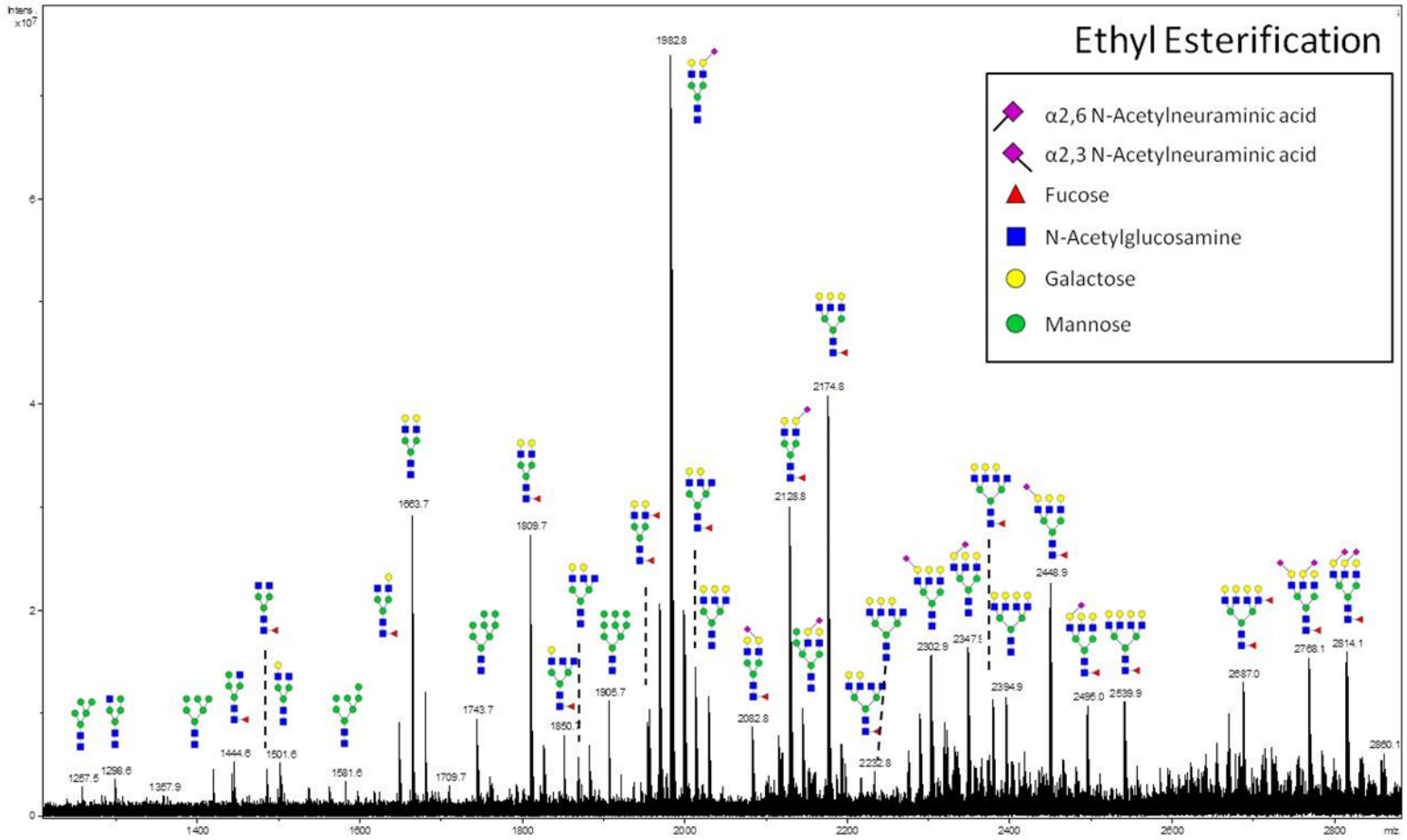


Figure 33. Off-Tissue Ethyl Esterification of HCC Glycans Correlates with MALDI-IMS Results. N-glycans extracted from HCC tissue sections were modified and enriched by ethyl esterification followed by cotton HILIC tips. Enriched glycans were spotted on a MALDI plate and analyzed in positive ion mode. An annotated spectrum with major glycoforms is provided.

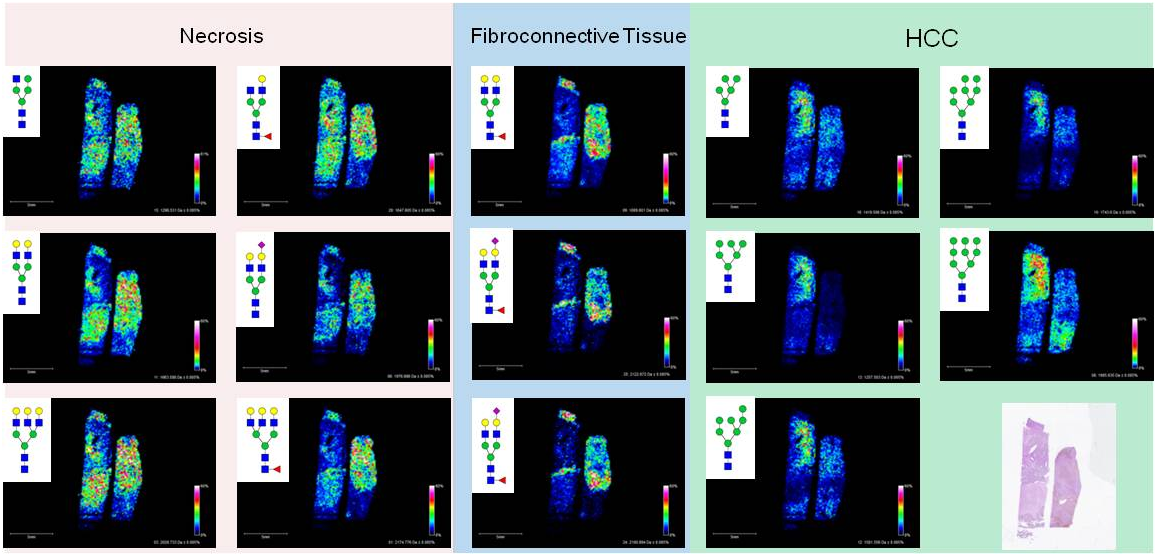


Figure 34. N-glycans Can Distinguish Complex Histopathology in HCC Tissue. N-glycan distribution mapped to specific histological regions in an HCC tissue. High mannose N-glycans were more abundant in the HCC tissue regions, while complex and hybrid N-glycans were more abundant in the necrotic and fibrous tissue regions.

Table 3. Comparative Analysis of N-glycans from MALDI-IMS and Ethyl Esterification.

		MALDI-IMS		Ethyl Esterification	
Observed Glycan	Database m/z	Observed	m/z	Observed	m/z
Hex5HexNAc2	1257.422	Yes	1257.497	Yes	1257.422
Hex4HexNAc3	1298.449	Yes	1298.522	Yes	1298.450
Hex6HexNAc2	1419.476	Yes	1419.530	Yes	1419.473
Hex4dHex1HexNAc3	1444.507	Yes	1444.546	Yes	1444.507
Hex3dHexHexNAc4	1485.534	Yes	1485.587	Yes	1485.540
Hex4HexNAc4	1501.529	Yes	1501.592	Yes	1501.536
Hex7HexNAc2	1581.528	Yes	1581.562	Yes	1581.532
Hex4dHex1HexNAc4	1647.587	Yes	1647.617	Yes	1647.589
Hex5HexNAc4	1663.581	Yes	1663.582	Yes	1663.591
Hex8HexNAc2	1743.581	Yes	1743.562	Yes	1743.592
Hex5dHex1HexNAc4	1809.639	Yes	1809.661	Yes	1809.644
Hex4dHex1HexNAc5	1850.666	Yes	1850.686	Yes	1850.683
Hex5HexNAc5	1866.661	Yes	1866.663	Yes	1866.686
Hex9HexNAc2	1905.634	Yes	1905.612	Yes	1905.647
Hex5HexNAc4NeuAc1	1954.677	Yes	1954.589 1976.707[e]	Yes	1936.692[a] 1982.733[b]
Hex5dHex2HexNAc4	1955.697	Yes	1955.692	Yes	1955.719
Hex5dHex1HexNAc5	2012.719	Yes	2012.711	Yes	2012.745
Hex6HexNAc5	2028.714	Yes	2028.737	Yes	2028.737
Hex5dHex1HexNAc4NeuAc1	2100.735	Yes	2100.759 2122.720[e]	Yes	2082.733[a] 2128.798[b]
Hex5dHex2HexNAc5	2158.777	Yes	2158.806	No	N/A
Hex6dHex1HexNAc5	2174.772	Yes	2174.768	Yes	2174.807
Hex5dHex1HexNAc6	2215.798	Yes	2215.847	Yes	2215.837
Hex6HexNAc6	2231.973	No	N/A	Yes	2231.828
Hex6HexNAc5NeuAc1	2319.809	Yes	2319.804 2341.804[e]	Yes	2301.869[a] 2347.878[b]
Hex6dHex2HexNAc5	2320.829	Yes	2320.746	Yes	2320.902
Hex6dHex1HexNAc6	2377.851	Yes	2377.859	Yes	2377.897
Hex7HexNAc6	2393.846	Yes	2393.796	Yes	2393.888
Hex6dHex1HexNAc5NeuAc1	2465.867	Yes	2465.803 2487.886[e]	Yes	2447.904[a] 2493.916[b]
Hex6dHex3HexNAc5	2466.887	Yes	2466.716	No	N/A
Hex7dHex1HexNAc6	2539.904	Yes	2539.856	Yes	2539.886
Hex7dHex2HexNAc6	2685.962	Yes	2686.010	Yes	2686.028
Hex6dHex1HexNAc5NeuAc2	2756.962	Yes	2756.976 2778.961[e]	Yes	2767.056[c] 2813.066[d]
Hex8HexNAc7	2758.978	Yes	2758.976	No	N/A

3.4.3.4 Glycosylation can Distinguish Cancer Subtypes

The ability of N-glycans to distinguish tissue subtypes and pathologies does not only pertain to HCC. A higher-throughput, TMA-based imaging experiment was performed to characterize and compare a large number of renal cancer samples in a single experiment. In the data provided, the classification of each set of tissue cores is outlined by a specified color. Examples of normal cortex (blue), normal medulla (purple), clear cell carcinoma (green), transitional cell carcinoma (red), chromophobe renal cell carcinoma (yellow), granular cell carcinoma (light blue), and carcinosarcoma tissue cores (white) are represented in the TMA. Hex5dHex2HexNAc5 (m/z = 2158.918) was elevated in all six normal cortex tissue cores (Figure 35a), while Hex5dHex1HexNAc4 (m/z = 1809.705) was consistently elevated in the normal medulla cores (Figure 35b). An overlay of Hex5dHex2HexNAc5 (blue) and Hex5dHex1HexNAc4 (purple), is capable of distinguishing normal kidney anatomy (Figure 35f). The expression of other N-glycans varied across tumor tissue cores with various specificity. In comparison to all other tissue subtypes, Hex5dHex2HexNAc4 (m/z = 1955.748) was elevated the most in clear cell carcinoma tissues (Figure 35d), while Hex7HexNAc2 (m/z = 1581.569) was most abundant in the granular cell carcinoma example (Figure 35c). In an overlay of Hex5dHex2HexNAc5 (blue), Hex5dHex1HexNAc4 (purple), and Hex5dHex2HexNAc4 (green), the clear cell carcinoma tissue cores are distinguished from the normal cortex and medulla cores (Figure 35g). Finally, there is an enormous degree of heterogeneity across the tumor cores, an example being Hex5HexNAc4 (m/z = 1663.570) which is much more abundant in two of the clear cell carcinoma cores in comparison to the rest of the clear cell carcinoma cores (Figure 35e). The glycans selected for display are representative of a larger population of N-glycans that were observed in the experiment. Interestingly, N-glycans in the transitional cell carcinoma were consistently lower in abundance than the other tissue cores on the TMA.

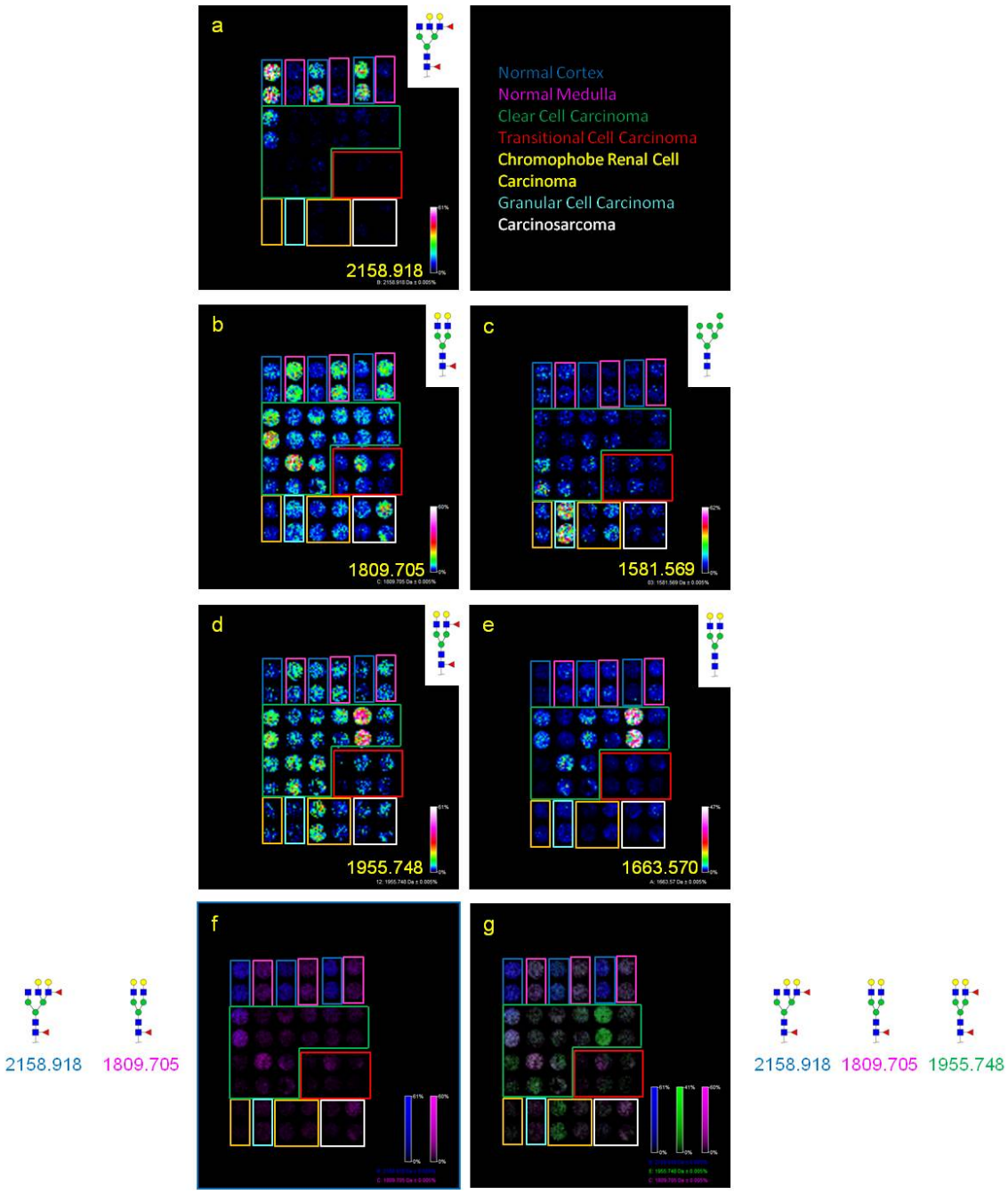


Figure 35. Aberrant Glycosylation in RCC TMA can Distinguish Cancer Subtypes. A RCC TMA was processed by MALDI imaging and revealed individual glycans or multiple glycans that offer diagnostic potential for disease or cancer subtypes. Hex5HexNAc5dHex2 is elevated in the normal cortex compared to all other tissue types (a), while Hex5HexNAc4dHex1 is more abundant in the normal medulla (b). Man7 was elevated in granual cell carcinoma (c) while Hex5HexNAc4dHex2 (d) and Hex5HexNAc4 (e) were elevated in the clear cell carcinoma cores. Furthermore, various overlays could differentiate the tissue subtypes (f-g).

3.4.4 **Discussion**

Profiling of N-glycans in fresh/frozen tissues, FFPE tissue blocks and TMAs is a new application of MALDI-IMS methodology that has the potential to identify systemic disease markers and elucidate disease etiology. This methodology is particularly relevant for cancer tissues, as most known cancer biomarkers are glycoproteins or carbohydrate antigens, including AFP for HCC (219,434,451–457). The MALDI-IMS glycan profiling data presented herein, although from only a limited number of HCC tissues, illustrates this potential. Compared to other biomolecules targeted by MALDI-IMS like lipids, metabolites and proteins/peptides, N-glycans offer a particular advantage in that there are relatively fewer total glycan ion signals that are detected in a given tissue. Assigning the underlying composition for each glycan is rather straight-forward on the basis of accurate masses, albeit anomeric linkages are not differentiated. In addition to mass accuracy, glycan compositions can be validated by a variety of approaches, such as CID and derivatization. Structural standards and several glycan reference databases are available (407,459).

Knowing the identity of the glycan offers many other opportunities for application to cell biology and genomic studies, particularly as these structures are synthesized by specific glycosyltransferases, and modified by specific glycosidases. This facilitates comparative analysis of genomic data for these enzymes from the same tissues. Knowing the glycan composition also allows inference of different cell biology aspects related to biosynthesis, vesicular trafficking or degradation processes involving N-glycans. We have found across multiple tumor types, and illustrated herein, that identifying a specific tumor glycan profile in tissues is fairly straightforward (459). Herein, this principle has been further expanded as consistent glycan signatures for fibroconnective tissue have been observed across cancer types (Figure 31). We are further testing whether detecting changes in glycan profiles can be used as surrogate indicators of

changes in glycoprotein expression associated with cancer development and progression, as well as responses to treatment. Current efforts are also directed at using the N-glycan profiles as a tissue map to identify disease-specific regions of interest for subsequent glycopeptide analysis strategies.

There still remain many challenges associated with improving the approach, some of which were also highlighted herein. Detection of high mass N-glycans by MALDI mass spectrometry, particularly those glycans containing sialic acids, has always been a challenge as previously described (230,248,249). We are addressing these challenges in three ways, at the instrumentation level, with on-tissue analysis approaches, and with off-tissue analysis approaches. The source configuration of the SolariX 7T FTICR instrument used for these experiments is already inherently amenable to better detection of sialic acid glycans. Compared to most MALDI-TOF instruments (non-FTICR), less vacuum and a cooling gas following laser desorption/ionization is present. This combination has been reported to aid in the retention of sialic acids on glycoconjugates. The ability to selectively capture specific glycan ions directly from tissue for CID is also an advantage of the SolariX instrument (459). Comparison of N-glycan profiles of the same tissue slides in different MALDI-TOF instruments is ongoing, with emphasis on how detection of sialic acid containing glycans differ across instrument configurations. In this context, different MALDI matrix formulations are also being tested and compared. These are all cumulatively being assessed in the optimization for detection of higher mass glycan species.

Based on mass accuracy of standards and derivatization experiments, we show that while sialylated N-glycans are detected in MALDI-IMS (Figure 30a), they appear to be of lower abundance than may be naturally occurring (Figure 33) (431,459). To correct for this limitation, different on-tissue and off-tissue derivatization approaches of the N-glycans are being investigated. Attempts to globally modify N-glycan reducing ends (i.e. 3-aminoquinoline) or

specifically modifying labile sialic acid residues (i.e. p-toluidine or ethyl esterification) represent potential solutions (235,237). We believe that some modification of the above off-tissue derivation approaches used for N-glycans could be adapted to an on-tissue MALDI-IMS profiling workflow. Additionally, performing MALDI-IMS in negative ion mode offers another reasonable option for partially retaining sialic acids of N-glycans (247).

Given the ability to use FFPE tissues as the primary source for N-glycan imaging and their general availability, there is a potential to leverage this approach to discover glycan biomarkers of HCC and other cancer types. To facilitate biomarker discovery, the described techniques must be used with large numbers of tissue samples on sample slides or in tissue microarray format, highlighting the need for an accurate method to directly compare between samples/experiments. In the example provided within, N-glycan signatures were capable of distinguishing both tumor from non-tumor tissue cores, but also between tumor subtypes (Figure 35). To determine the significance of these glycans, a similar study involving an increased number of tissue cores is required. Ongoing research is directed at determining best practices for data normalization and incorporating internal standards into experimental workflows. These issues are common to all MALDI-IMS experiments (460,461). One apparent solution is to utilize a biologically unique glycan as an internal standard that can serve dual purpose to calibrate in the appropriate mass range as well as serve to normalize signal between samples/experiments. By incorporating such an internal standard into the matrix spray, it would allow to correct for differences in matrix deposition, tissue ionization, as well as changes in signal intensity across experiments. Likewise an approach using informative peaks may be the most appropriate method. Only after large enough data sets are generated will we be able to determine the best approach to analyze these larger sample sets. Use of TMAs for glycan MALDI-IMS can certainly facilitate the increase in sample numbers (459).

In conclusion, the profiling of N-glycans offers many new avenues of research that can be initiated by, and complemented with, MALDI-IMS strategies. Challenges and limitations of the approach are evident, but we believe most can be addressed moving forward as more samples are analyzed, new derivation and matrix strategies tested, and long-term, continued improvement in instrumentation platforms.

Chapter 4: Application of Novel Glycomic Approaches to Identify Pancreatic Cancer Disease Markers

4.1 Introduction

Pancreatic cancer represents one of the most deadly cancers, both in terms cancer related deaths per year and 5-year survival rates. The dismal 5-year survival rate following detection, which is around 6%, can be attributed to poor understanding of disease etiology, rapid disease progression, late diagnosis, limited effective treatment options, and resistance to therapeutic intervention (2). Patient outcome is significantly improved when the cancer is detected at an early stage and before it has spread to regional lymph nodes or metastasized to distant sites. This is largely because these patients are eligible for surgical resection, which represents the only curative treatment for pancreatic cancer (3,5,8). Unfortunately, it is estimated that only 9% of patients are diagnosed with localized pancreatic cancer (5). Therefore, efforts directed at the early diagnosis of pancreatic cancer and further understanding of biological pathways implicated in disease progression are extremely important.

An in-depth analysis of aberrant glycosylation, a hallmark of all cancers, offers the potential to advance the understanding of pancreatic cancer and identify new disease markers. Furthermore, glycoproteins and glycan antigens are ideal targets for initial discovery phase disease marker studies, as they [1] are prevalent in biological fluids, [2] play a direct role in cancer progression and metastasis, and [3] represent a majority of the protein-based FDA approved biomarkers. Both individual glycan antigens and alterations in the glycosylation biosynthetic machinery have been implicated in pancreatic cancer. The carbohydrate antigen 19-9 (SLeA, CA19-9) is the most well defined disease marker for pancreatic cancer and the target of the only FDA approved blood test for the management of pancreatic cancer (332,333). Functionally, CA19-9 has been implicated in disease metastasis through the interaction of the antigen with endothelium E-selectin (337). However, CA19-9 is not used as a clinical disease marker for pancreatic cancer, as elevated levels are detected in benign diseases and other sites of cancer (341–343). Additionally, between 5-15% of the population do not express

fucosyltransferase 3 and are therefore incapable of expressing CA19-9, even with advanced pancreatic cancer (333,343). While elevated CA19-9 is the most pronounced glycosylation-related change in pancreatic cancer, global alterations in glycosylation have been detected. At a fundamental level, the malignant processes resulting in the formation of pancreatic cancer involve several common mutations in oncogenes and tumor suppressors. These mutations ultimately affect the glycan biosynthesis pathway in pancreatic cancer. Constitutive KRAS activation increases the expression of GLUT1, HK1, HK2, PFK1 and LDHA, all of which impact glycosylation by directing glycolytic intermediates to the HBP (40,361). Similarly, the dense stroma and ECM accumulation surrounding pancreatic cancer results in a hypoxic microenvironment and initiates anaerobic glucose metabolism through glycolysis (354,355). The shift to glycolysis is accompanied by an upregulation of GLUT1, HK2, PFK1, PFK2, PKM2, GFPT1 and GFPT2 (354,462). Importantly, in vitro knockdown of GFPT1, a key modulator of entry into the HBP, exhibited the same reduction in tumor growth as the extinction of mutated KRAS, directly implicating glycosylation in tumor growth (40). In addition to enzymatic regulation of the hexosamine biosynthetic pathway, glycosyltransferases have been reported to be alternatively expressed in pancreatic cancer (336,463–465). Alterations in the hexosamine biosynthetic pathway and glycosyltransferase expression are apparent in global glycoform analysis, where increases in fucosylation, sialylation and glycan branching have been observed (343,373). However, the pancreatic cancer glycome has been analyzed on a relatively small subset of patients, often ignoring individual glycoforms in favor of structural motifs.

Recent analytical advances have improved the ability to monitor changes in the glycome in relation to cancer status. Traditional techniques characterized fluorescently labeled glycans based off of their retention time in chromatographic separations or by sequential enzymatic digestion (466). As chromatographic separation is required for each individual sample, this method is not ideal for high-throughput disease marker experiments. Other studies have utilized

mass spectrometers to characterize disease-related changes in glycan expression levels, although lengthy desalting and derivatization steps are frequently employed (467). When applied to the analysis of biological tissues, these methods result in a complete loss of glycan localization. As up to 90% of the tumor volume is occupied by surrounding stromal cells, this analysis may differentiate the non-diseased tissue from the stroma, as opposed to the tumor (24). Conversely, immunostaining with lectins or monoclonal antibodies directed at specific glycosylation structures or antigens can probe glycan localization within tissue sections, but does not target or identify individual glycoforms. To this end, our laboratory has introduced a novel approach to study specific N-linked glycan localization using MALDI-IMS on fresh frozen and FFPE tissue sections (431,459). This technique has been further explored and validated by other laboratories (468,469). In addition to simply localizing and identifying N-glycans in tissue sections, we have recently demonstrated the potential of the method for the identification of individual glycans that discriminate tumor from non-tumor tissue specimen in a high-throughput TMA-based approach.

The present aim is an example of how MALDI-IMS of N-linked glycans can be adapted to the identification of disease markers for pancreatic cancer. In the aim herein, we report MALDI-IMS of N-glycans in TMAs and whole tissue blocks, followed by subsequent glycan characterization by mass accuracy and derivatization. As the spatial localization of N-glycans is retained during MALDI-IMS, analysis of tissue blocks can be further interrogated to look at glycosylation patterns in histological regions other than the tumor and non-tumor regions. Finally, N-glycan distribution was correlated with CA19-9 staining on serial slides, to determine any correlations between CA19-9 and released N-glycans.

4.2 Materials and Methods

4.2.1 Materials

Trifluoroacetic acid, α-cyano-4-hydroxycinnamic acid (CHCA), sodium hydroxide (NaOH), and 1-hydroxybenzotriazole hydrate (HOBt) were obtained from Sigma-Aldrich (St. Louis, MO). HPLC grade methanol, ethanol, acetonitrile, xylene and water were obtained from Fisher Scientific (Pittsburgh, PA). 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) was obtained from Oakwood Chemical (West Columbia, SC). Tissue Tack microscope slides were purchased from Polysciences, Inc (Warrington, PA). Citraconic anhydride for antigen retrieval was from Thermo Scientific (Bellefonte, PA). Recombinant Peptide N-Glycosidase F (PNGaseF) from *Flavobacterium meningosepticum* was expressed and purified as previously described, and is available commercially as PNGase F Prime™ from Bulldog Bio (Portsmouth, NH) (431). Cotton tips for HILIC enrichment of N-glycans were produced using 100% cotton swabs from Assured (Rio Ranch, NM).

4.2.2 FFPE Tissues, TMAs, and Plasma

All human tissues and TMAs used were de-identified and determined to be not human research classifications by the respective Institutional Review Boards at MUSC. For each section analyzed, histological analysis and staining with hematoxylin and eosin (H&E) was performed and regions of diverse pathology were annotated by a pathologist. Plasma from patients with pancreatitis and PDAC were de-identified and determined to be not human research classifications by the respective Institutional Review Boards at MUSC.

4.2.3 Sample Preparation for MALDI Imaging

Tissues and TMA blocks were sectioned at 5μm and mounted on slides (25 x 75mm) compatible with the Bruker slide adaptor. After mounting, sections were dewaxed and antigen retrieval proceeded as described in Section 3.3.3.3.

PNGaseF (20μg/slide) was applied to slide-mounted tissue blocks and TMAs using the ImagePrep spray station (Bruker Daltonics) as previously described (459). N-glycan release occurred during a 2hr incubation at 37°C in a humidified chamber, followed by desiccation and matrix application. For the TMAs, α-Cyano-4-hydroxycinnamic acid matrix (CHCA), consisting of 0.021g CHCA in 3mL 50% acetonitrile/50% water and 12μL 25%TFA, was applied using the ImagePrep sprayer. The TM-Sprayer (HTX Imaging) was used to coat slides containing whole tissue blocks with CHCA. CHCA was prepared at a concentration of 5mg/mL in 50% ACN/50% H_2O at .1% TFA. CHCA was applied at 70°C at .0017mg/mm^2.

4.2.4 Glycan Derivatization

N-glycans were extracted from slides as described previously and dried by vacuum centrifugation (431). The ethyl esterification protocol, including the modification and enrichment, was adapted from Reiding et al. (237). The protocol proceeded as described in Section 3.4.2.5.

4.2.5 CA19-9 and SLeX Staining

Serial sections of five of the TMAs were stained with CA19-9 and SLeX. Individual cores were manually scored for the presence or absence of stain and annotated for the presence of tumor or non-tumor tissue. Sensitivity, specificity, positive predictive value and negative predictive value were calculated for both stains individually across the five TMAs.

4.2.6 Data Processing

Imaging data was loaded into FlexImaging 4.1 (Bruker Daltonics) for visual analysis of tissue blocks and TMAs. TMAs were searched against an N-glycan library to identify N-glycans present across all six TMAs. For computational peak picking of the TMAs, Regions of Interest (ROIs) representing each tissue core were exported from data analysis using the hierarchical clustering option and further processed using an in-house workflow.

To assess the mass accuracy of N-glycans, N-glycans present in all six TMAs were manually tabulated along with the accurate mass of the glycans. In addition to the accurate mass, m/z values were reported from the imaging peak picking tool as well as individual spectra from a tissue imaging experiment. To generate the m/z values from individual spectra, one spectra localized to the tumor tissue and one spectra localized to the non-tumor tissue were loaded into FTMS processing (Bruker Daltonics) and recalibrated. PPM errors were calculated for both the picked peaks and the individual spectra.

4.2.7 Glycomic Analysis of Plasma Pools

Plasma from patients with pancreatitis and PDAC were pooled together for glycomic analysis. 10µL of plasma from each pool was diluted in 90µL H_2O prior to digestion with 1µg PNGaseF overnight at 37°C. N-glycans were extracted from the sample following the addition of 400µL MeOH, 100µL $CHCl_3$, and 300µL H_2O and centrifugation at 14,000 x g for 2 minutes. The aqueous phase was collected, dried by vacuum centrifugation, and ethyl esterified as described in Section 3.4.2.5.

4.3 Results

4.3.1 N-Glycan Variation in Complex Histopathology Regions

FFPE tissue blocks containing both pancreatic cancer and matched non-cancer tissues, as well as other complex histopathology regions, were selected for N-glycome analysis by MALDI-IMS. N-glycans in these tissue sections displayed patterns of regionalized distribution that correlated with histopathology analysis by H&E staining. In the example provided, the regions of the tissue section are outlined in different colors, with each color corresponding to a different histological region, as annotated by a pathologist (Figure 36a). These regions are pancreas tumor/ pre-cancerous lesions (green), intestine mucosa (yellow), fibroadipose connective tissue (blue), smooth muscle (orange), and non-tumor pancreas tissue (red). N-glycans were identified that

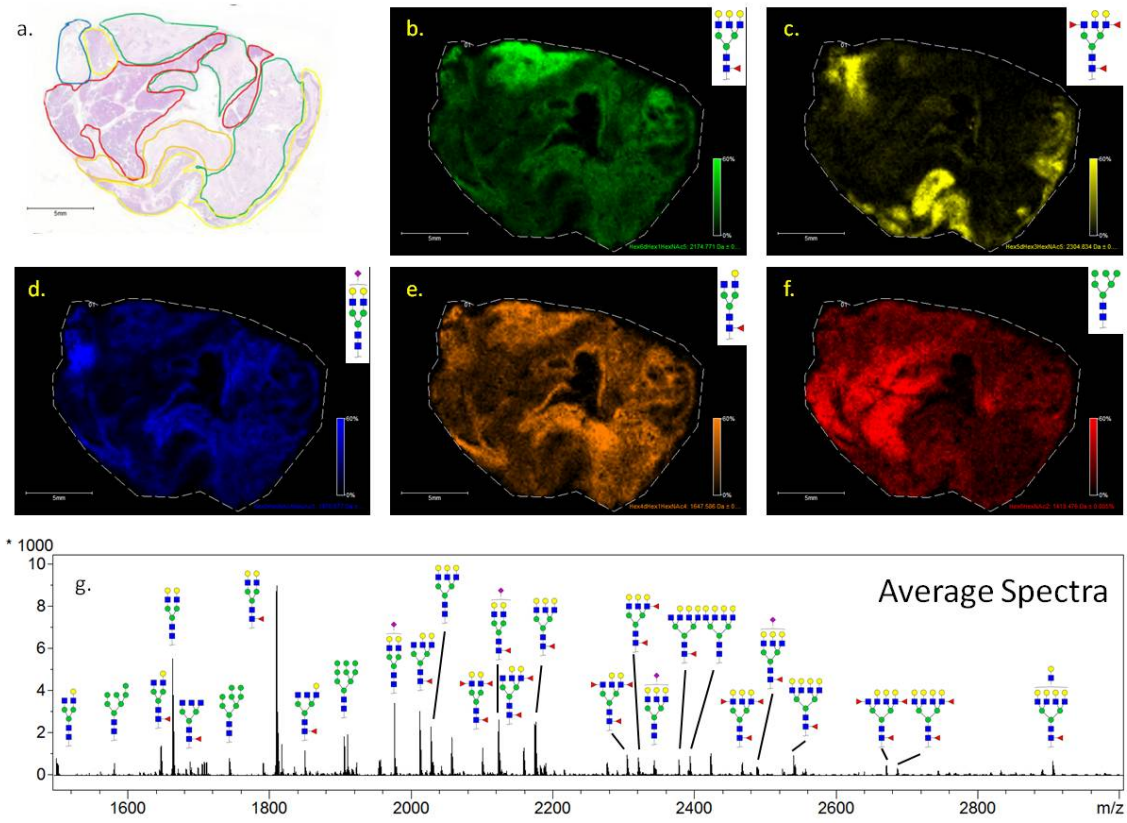


Figure 36. Heterogeneous Distribution of N-Glycans in Pancreatic Cancer Tissue Sections. A PDAC tissue section of complex histology was profiled by MALDI-IMS. In the H&E, tumor/ pre-cancerous lesions (green), intestine mucosa (yellow), fibroadipose connective tissue (blue), smooth muscle (orange), and non-tumor pancreas tissue (red) are all outlined. In the MALDI-IMS images, individual glycans correlate with defined histology regions. Hex6dHex1HexNAc5 (b, tumor), Hex5dHex3HexNAc5 (c, intestine mucosa), Hex5HexNAc4NeuAc1 (d, fibroadipose connective tissue), Hex4dHex1HexNAc4 (e, smooth muscle), and Hex6HexNAc2 (f, non-tumor pancreas) are among these N-glycans.

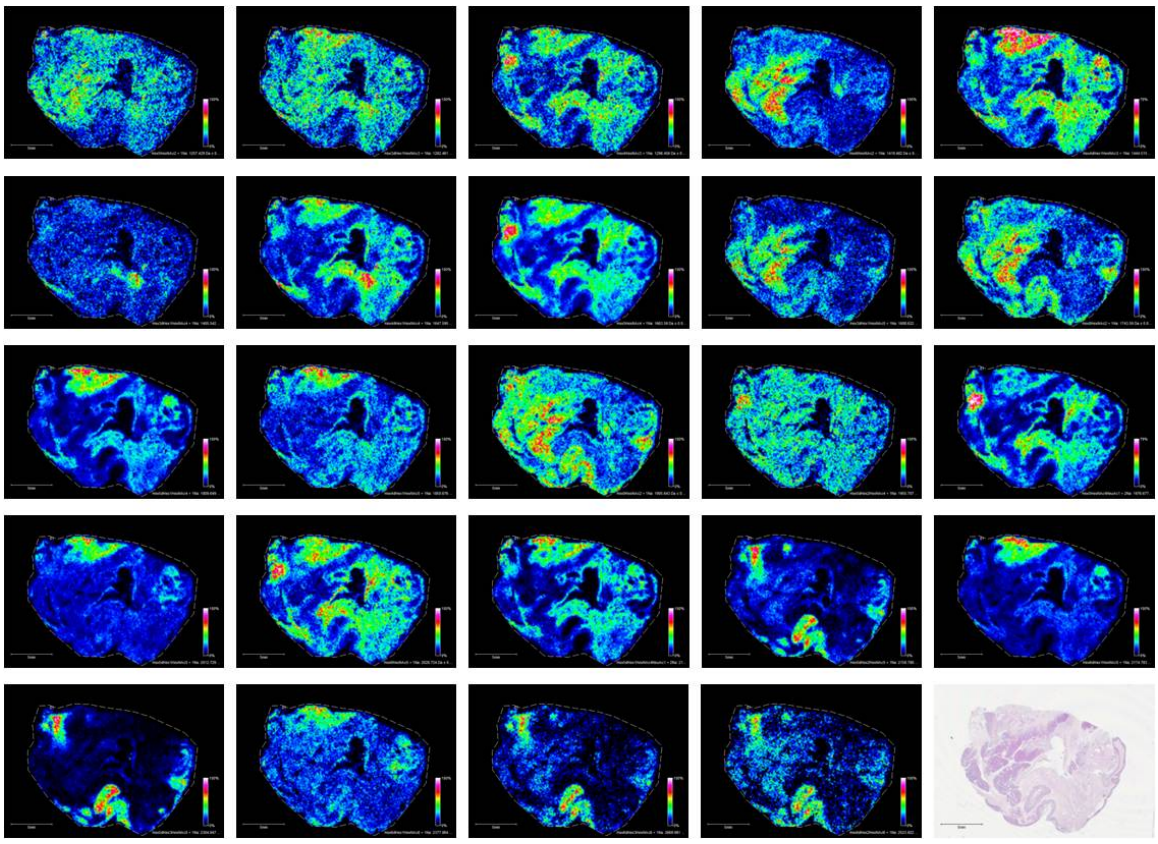


Figure 37. Panel of N-Glycan Distribution in PDAC Tissue Section. Upwards of 90 ions corresponding to N-glycoforms were observed in an individual tissue section. In a representative panel of these glycoforms, many of these glycans exhibit a distribution pattern related to regions of complex histology.

were either highly elevated or nearly exclusively detected in each region. Hex6dHex1HexNAc5 (m/z = 2174.771) is elevated in the tumor region of the pancreas, which contrasts with Hex6HexNAc2 (m/z = 1419.476), which is elevated non-tumor pancreas (Figure 36b,f). Similarly, Hex5HexNAc4NeuAc1 (m/z = 1976.677) and Hex4dHex1HexNAc4 (m/z = 1647.586) are elevated in the fibroadipose connective tissue and smooth muscle regions, respectively (Figure 36d-e). Unlike the other N-glycans above which are elevated in defined tissue regions but expressed at lower levels throughout the tissue section, Hex5dHex2HexNAc5 (m/z = 2304.834) is almost exclusively observed in the intestine mucosa (Figure 36c). Overall, close to 90 N-glycans were identified in the tissue section that correspond to our glycan library, most having a distribution that correlates with a certain histology-derived region of the pancreas. The most abundant N-glycans are annotated in the overall average spectra from the imaging experiment (Figure 36g). A selected panel of N-glycans demonstrates the differential distribution of N-glycans across the tissue section (Figure 37).

4.3.2 Individual N-Glycan Discriminators for Pancreatic Cancer

As N-glycan distribution varies between tumor and non-tumor regions of the pancreas in whole tissue blocks, we wanted to determine if individual glycans were consistently altered between tumor and non-tumor tissue sections across different patients. To this end, a higher-throughput study based on the analysis of six TMAs was performed to allow a larger number of samples to be surveyed in a reduced amount of time. Each TMA consisted of patient matched tumor and non-tumor tissue cores. Figure 38 provides a detailed workflow schematic of the data processing steps utilized to generate individual or panels of N-glycans that can distinguish tumor from non-tumor tissue cores. Initially, the data were loaded into an in-house peak picking algorithm, which identified 54 monoisotopic peaks present across all six TMAs. A total of 23 peaks that corresponded to N-glycans were identified from the 54 peaks (Table 4). The

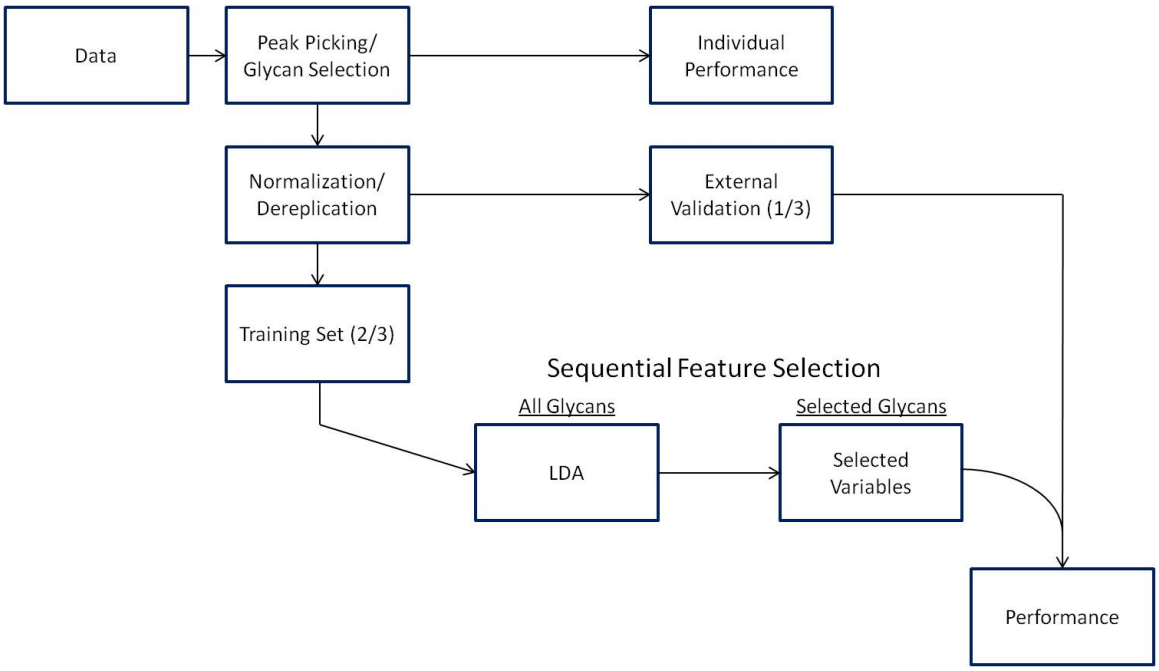


Figure 38. Workflow Schematic for the Identification of Individual and Panels of N-Glycan Disease Markers. Six TMAs were imaged by MALDI-IMS to add an element of throughput to the disease marker discovery process. Many glycoforms were detected in all 6 of the TMAs and were selected for further analysis of individual discriminators and as panels of biomarkers. For panel identification, 2/3rds of the data was used to optimize the variables for an LDA model, while the remaining 1/3rd of the data was used to test the performance of the model.

Table 4. Individual Glycan Discriminators for Pancreatic Cancer.

m/z bin	Glycan Structure	log_2FC (T/N)	BH p-value
1419.4818	Hex6HexNAc2 + 1Na	-0.08	0.8223
1442.5168	Hex4dHex1HexNAc3 + 1Na	0.90	<0.0001
1485.5341	Hex3dHex1HexNAc4 + 1Na	0.69	0.0005
1501.5379	Hex4HexNAc4 + 1Na	0.91	<0.0001
1581.5024	Hex7HexNAc2 + 1Na	0.21	0.1116
1647.5518	Hex4dHex1HexNAc4 + 1Na	0.82	<0.0001
1688.5851	Hex3dHex1HexNAc5 + 1Na	0.78	0.4828
1743.5649	Hex8HexNAc2 + 1Na	-0.14	0.3706
1809.5829	Hex5dHex1HexNAc4 + 1Na	1.12	<0.0001
1850.6041	Hex4dHex1HexNAc5 + 1Na	1.18	<0.0001
1954.5931	Hex5HexNAc4NeuAc1 + 1Na	0.54	0.1175
1976.6138	Hex5HexNAc4NeuAc1 + 2Na	0.47	0.0119
2012.7156	Hex5dHex1HexNAc5 + 1Na	1.11	<0.0001
2028.6388	Hex6HexNAc5 + 1Na	0.91	<0.0001
2100.6539	Hex5dHex1HexNAc4NeuAc1 + 1Na	0.75	0.0002
2122.6127	Hex5dHex1HexNAc4NeuAc1 + 2Na	0.75	<0.0001
2158.6946	Hex5dHex2HexNAc5 + 1Na	0.42	0.0175
2174.773	Hex6dHex1HexNAc5 + 1Na	1.11	<0.0001
2320.7524	Hex6dHex2HexNAc5 + 1Na	-0.05	0.8318
2393.8019	Hex7HexNAc6 + 1Na	0.93	<0.0001
2539.8632	Hex7dHex1HexNAc6 + 1Na	0.88	<0.0001
2742.928	Hex7dHex1HexNAc7 + 1Na	2.29	<0.0001
3270.0758	Hex9dHex1HexNAc8 + 1Na	0.94	<0.0001

expression level of the glycans was assessed by their p-value and Log_2FC (Table 4). All N-glycans were detected as the $[M + Na]^+$ adduct, except for sialylated glycans which were observed as singly or doubly sodiated ions. Of the peaks corresponding to glycans, 17 were elevated in the tumor tissues at a statistically significant level ($p \leq .05$), while none were elevated in the non-tumor cores at a statistically significant level. Initially peaks were characterized as N-glycans by comparing the observed m/z values to a library of N-glycan structures, but peaks were further validated by glycan release and derivatization using ethyl esterification of sialic acids. For this approach, glycans were extracted from a tissue section, underwent derivatization and enrichment, and were spotted on a MALDI plate prior to analysis. Of the 23 peaks corresponding to N-glycans in the structural library, all 23 were observed in the glycan derivatization experiment (Sup Sheet 5). However, additional N-glycans were detected in the analysis following derivatization. The five N-glycans with the highest Log_2FC values were all complex, fucosylated N-glycans, and all had a $Log_2FC \geq 1.11$. While the level of these glycans were significantly different between tumor and non-tumor ($p \leq .05$), individual discriminators displayed poor sensitivity and specificity for distinguishing the conditions. For example, Hex7dHex1HexNAc7 (m/z = 2742.983) (data not shown) was only elevated in a small subset of tumor tissues, while Hex6dHex1HexNAc5 (m/z = 2174.771) was elevated in a majority of tumor cores but also some non-tumor cores (Figure 39e). In representative images from five pancreas tissue blocks, Hex6dHex1HexNAc5 was primarily localized to the tumor tissue, but was frequently detected in regions not characterized as tumor tissue by a pathologist (Figure 39a). This general trend is reflected in the statistical data, in which Hex6dHex1HexNAc5 is elevated in the tumor over the non-tumor tissue cores, with a $log_2FC = 1.11$ (Figure 39a). In contrast, m/z = 1419.476 (Hex6HexNAc2) is more ubiquitously expressed throughout the tissue blocks (Figure 39b) and TMA (Figure 39f), which is reflective of the low log_2FC of the Hex6HexNAc2 glycan, where $log_2FC = -.08$. Of the three high-mannose N-glycans detected across all six TMAs, all had a low

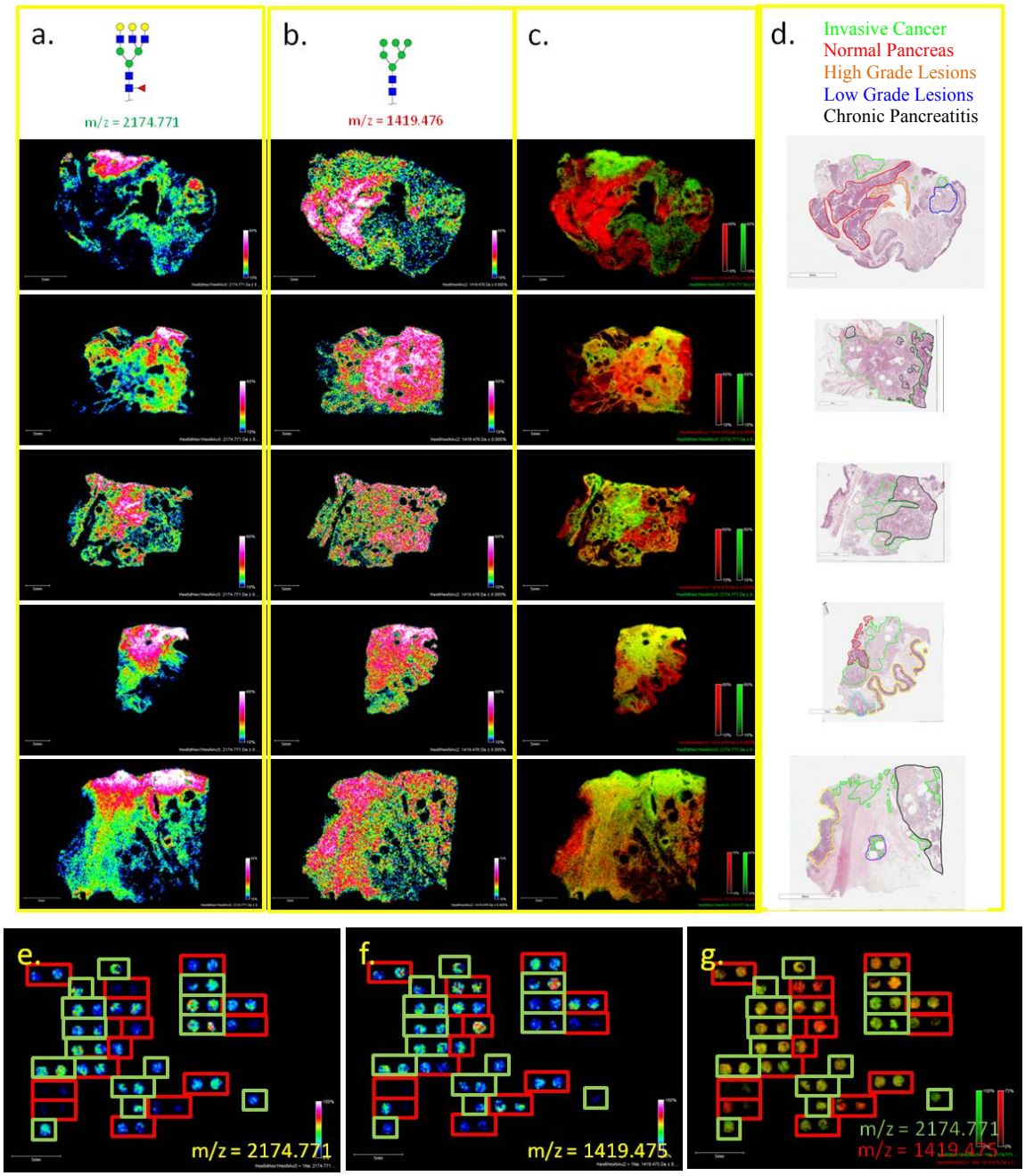


Figure 39. Combinations of Individual Discriminators Reveal More Robust Differences in Tumor and Non-Tumor Samples. Representative images of 5 tissue sections of complex histology and 2 tissue microarrays, looking at the localization of Hex6dHex1HexNAc5 (a,e), Hex6HexNAc2 (b,f), and the overlay of the two glycans (c,g). In the TMA, tumor cores are outlined in green, while non-tumor cores are outlined in red. Furthermore, in overlay images, Hex6dHex1HexNAc5 is presented in green, while Hex6HexNAc2 is presented in red. The overlay of the two glycans is able to distinguish tumor from non-tumor cores in both the whole tissue blocks and the tissue microarray.

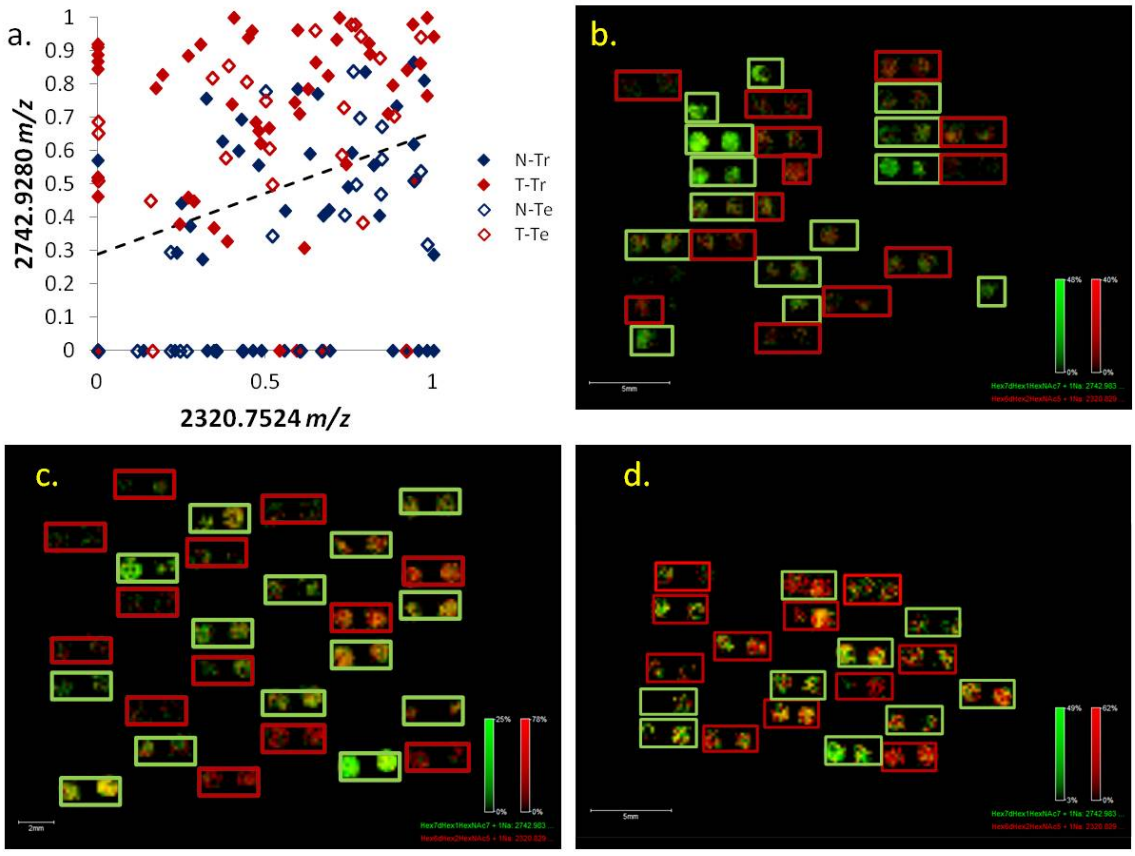
or negative Log_2FC value and were not significantly different between the tumor and non-tumor cores.

4.3.3 N-Glycan Biomarker Panels

The combination of N-glycans improved the ability of the platform to distinguish tumor from non-tumor tissues, both visually in the MALDI-IMS data and empirically in the statistical data. In comparison to the individual discriminators, the visual image overlay of Hex6dHex1HexNAc5 (green) and Hex6HexNAc2 (red) in both the tissue blocks (Figure 39c) and TMA (Figure 39g), yields more pronounced differences between the tumor from non-tumor tissue regions, as outlined by the pathologist (Figure 39d). To generate a N-glycan biomarker panel, a supervised machine learning algorithm was applied across the glycans identified in all six TMAs, focusing on the ability of the multiple N-glycans to distinguish tumor from non-tumor tissue cores (Table 4). With this approach, $2/3^{rd}$ of the data was subjected to a sequential feature selection process to minimize the misclassification error of the LDA model, and the remaining $1/3^{rd}$ was used as a validation set. In the LDA model, Hex7dHex1HexNAc7 (m/z = 2742.9280) and Hex6dHex2HexNAc5 (m/z = 2320.7524) displayed the lowest error rate of .1458 (Figure 40a). Additionally, the image overlay of Hex7dHex1HexNAc7 (green) and Hex6dHex2HexNAc5 (red) in three TMAs validates the statistics (Figure 40b-d). The predictive characteristics of the model were compared against CA19-9 and SLeX staining of the same TMAs. The LDA model consisting of two glycans was able to outperform both CA19-9 and SLeX staining in negative predictive value, and CA19-9 in positive predictive value (Table 5).

4.3.4 Comparison of Plasma and Tissue N-Glycome

Pooled plasma samples from patients with pancreatic cancer and pancreatitis were analyzed for differences in the N-glycome following ethyl esterification. Very few differences in the plasma N-glycome were observed in the two conditions (Figure 41a,b). This may be attributed

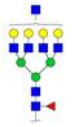
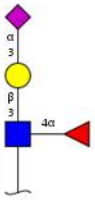
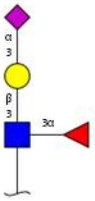
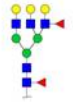


LDA

Correct Rate: 0.8542
Error Rate: 0.1458
Last Correct Rate: 0.8542
Last Error Rate: 0.1458
Sensitivity: 0.8696
Specificity: 0.8400
Positive Predictive Value: 0.8333
Negative Predictive Value: 0.8750
Positive Likelihood: 5.4348
Negative Likelihood: 0.1553

Figure 40. LDA Model of N-Glycan Discriminates Tumor from Non-Tumor Tissue Cores. Supervised machine learning algorithms, specifically the Linear Discriminate Analysis (LDA), were used to identify important features to distinguish tumor from non-tumor tissue sections. Two glycans (Hex7dHex1HexNAc7, m/z = 2742.9820; Hex6dHex2HexNAc5, m/z = 2320.7524) were observed to be important, and the performance of the LDA model based off of these two features was tested for performance of the biomarker panel (a). Representative images from three of the TMAs are presented with an overlay of Hex7dHex1HexNAc7 (green) and Hex6dHex2HexNAc5 (red) (b-d).

Table 5. Comparison of LDA Model to Carbohydrate Antigen Staining. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated based off of the markers ability to distinguish tumor from non-tumor tissue sections.

	LDA	CA19-9	SLeX
Sensitivity	0.8696	0.7321	0.8142
Specificity	0.8400	0.6139	0.8544
PPV	0.8333	0.6777	0.8598
NPV	0.8750	0.6739	0.8073
	 2742.9280		
	 2320.7524		

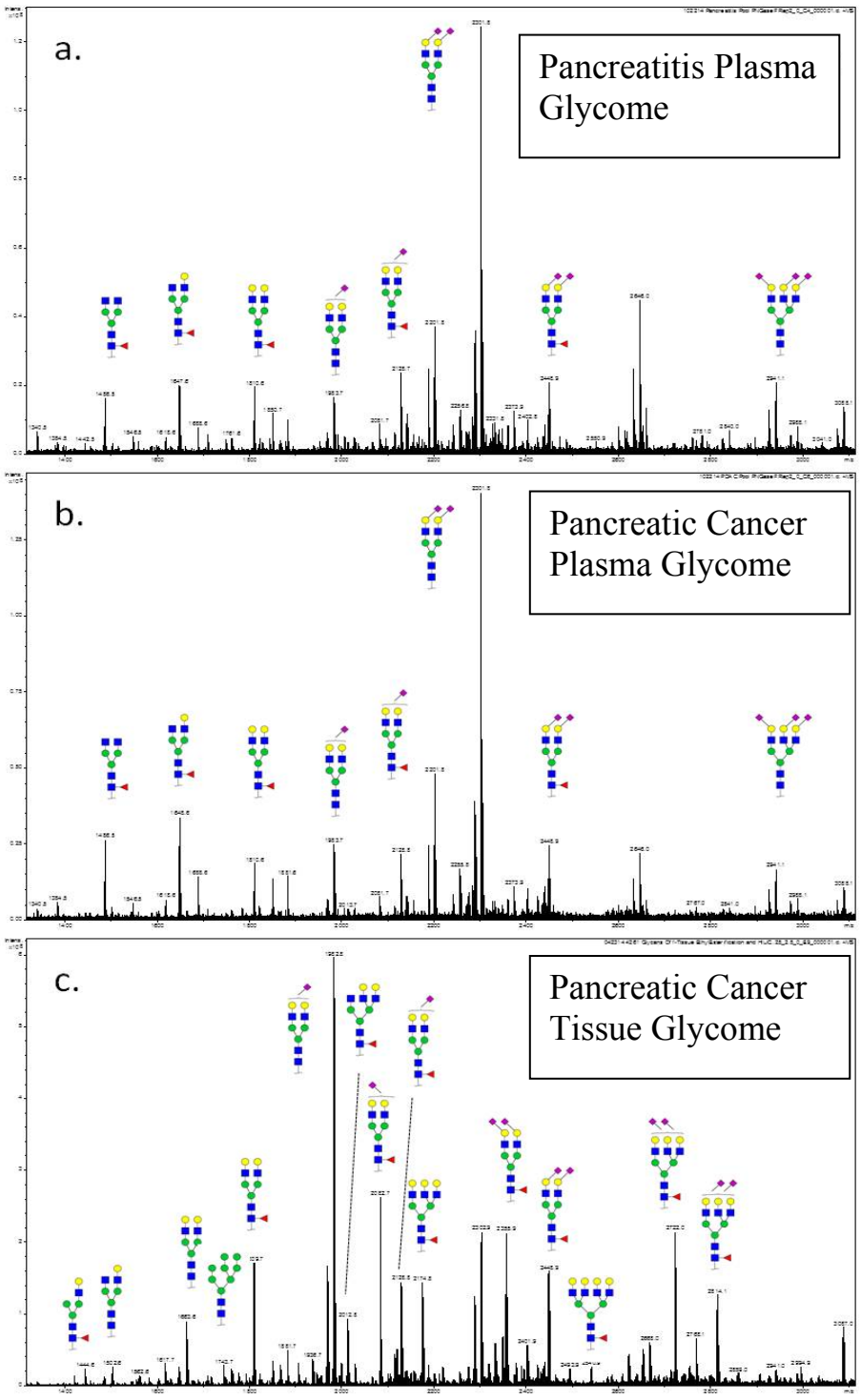


Figure 41. Pancreatic Cancer and Pancreatitis Plasma Glycome. The derivatized N-glycome from pooled plasma from patients with pancreatic cancer (b) and pancreatitis (a) exhibited little difference from one another. In contrast the pancreatic tissue glycome was more robust, suggesting the plasma glycome is not reflective of the tissue glycome.

to the presence glycoforms from abundant serum glycoproteins, such as immunoglobulins. On the other hand, the N-glycome from glycans extracted from a pancreatic cancer tissue block exhibited robust differences in the glycan species identified compared to the pancreatic cancer plasma pool (Figure 41c).

4.4 Discussion

The aim of this study was to investigate the ability MALDI-IMS analysis of N-glycans to be utilized as a disease marker identification platform for pancreatic cancer. To achieve this goal, TMAs were used for initial disease marker identification, but results were validated in both TMAs and whole tissue blocks. A secondary goal was to compare the classification metrics of individual discriminators to panels of N-glycans. Due to the enormous biological heterogeneity between individuals with cancer, single markers rarely have the sensitivity and specificity needed for screening asymptomatic patients. Biomarker panels, even when some of the markers are not adequate predictors of disease on their own, can improve disease classification (470). While aberrant glycosylation is known to be a hallmark of cancer, including pancreatic cancer, few studies have assessed alterations of N-linked glycosylation in pancreatic tissue sections. Instead, most studies monitor CA19-9 serum expression, which is primarily present on O-glycans and glycolipids (334,335). However, CA19-9 expression has well defined limitations for its use as a diagnostic disease marker. In our initial analysis of N-glycan distribution of pancreatic cancer tissue blocks, the expression levels of various N-glycans seem to correlate to complex histological regions present in the tissue blocks (Figure 36). The need for diagnostic markers for pancreatic cancer, limitations in currently available disease markers, and the observation of N-glycans that correlate to tumor and non-tumor tissue regions make pancreatic cancer an ideal disease model for this manuscript.

Given the lack of software available to process data from the SolariX FTICR mass spectrometer, mainly due to the enormous size of the high-resolution data, our laboratory

developed an in-house approach for peak picking and generation of biomarker panels (Figure 38). For the peak picking function, a bucket table consisting of m/z values and the corresponding relative intensity was generated by using the hierarchical clustering function present in the FlexImaging software. In the peak picking process, a total of 54 peaks were detected in all six TMAs, 23 of which correspond to masses present in our N-glycans structural library (Table 4, Sup Sheet2). Of the identified N-glycans, 20 are elevated in the tumor over the non-tumor tissue cores, 17 of which have a p-value $\leq .05$. The high-mannose glycans observed have a negative (Hex6HexNAc2, Hex8HexNAc2) or low (Hex7HexNAc2) $\log_2$FC value, with p-values $> .05$. It is clear that there is room for optimization in the peak picking process, as manual data processing identified peaks corresponding to 50 N-glycans that are present in all six TMAs (Sup Sheet 5). Furthermore, the peak picking selection, which uses binned windows for peak selection, exhibits higher m/z errors than individual processed spectra loaded into the FTMS processing software. After data processing, nearly all errors are below 10 PPM, which is significantly better than the PPM errors observed using the binned windows (Sup Sheet 5). Of the 23 peaks that correspond to N-glycans, all but one are detected in both the individual spectra and in the off-tissue derivatization and enrichment of N-glycans. As demonstrated by the table, many glycans are not observed in the peak picking algorithm but can be detected by FTMS processing and following glycan derivatization and enrichment (Sup Sheet 5).

Following the peak picking process, combinations of N-glycans were evaluated for their potential to distinguish tumor from non-tumor tissue sections, both visually and computationally. An overlay of two identified N-glycans, Hex6dHex1HexNAc5 and Hex6HexNAc2, is able to visually differentiate tumor from non-tumor tissues in whole tissue blocks and TMAs (Figure 39). For panel identification, the 23 peaks corresponding to N-glycans were subjected to supervised machine learning algorithms and sequential feature selection to identify N-glycan panels that are capable of distinguishing tumor from non-tumor tissue cores. The LDA model highlights the

importance of Hex7dHex1HexNAc7 and Hex6dHex2HexNAc5 in distinguishing tumor from non-tumor sections (Figure 40). This model outperforms CA19-9 staining metrics on serial TMA sections (Table 5). Both comparisons mentioned above, Hex6dHex1HexNAc5 vs. Hex6HexNAc2 and Hex7dHex1HexNAc7 vs. Hex6dHex2HexNAc5, contrast N-glycans with a large Log_2FC and a p-value $\leq .05$ with N-glycans with a negative Log_2FC and p-value $\geq .05$ (Table 4). Additionally, as little correlation was found between N-glycans and the CA19-9 antigen, it is possible to incorporate CA19-9 expression metrics into the biomarker panel to potentially improve the predictive value (470).

By validating the glycan distribution on whole tissue blocks, we were able to evaluate glycosylation trends in relation to complex histopathology. Among the findings, two important observations warrant further discussion. As previously mentioned, pancreatic cancer is only curable if detected at an early stage prior to becoming locally invasive or metastatic. As such, markers capable of detecting of pre-cancerous lesions are particularly useful. The overlay of Hex6dHex1HexNAc5 and Hex6HexNAc2 (Figure 39c, top tissue), is able to differentiate both low and high grade IPMN lesions from the non-tumor pancreas tissue. Interestingly, Hex6dHex1HexNAc5 (Figure 39a, top tissue) has a higher relative intensity in the tumor regions than the pre-cancerous lesions, suggesting a signature gradient that coincides with cancer stage. However, not all of these lesions ultimately become cancerous, so additional tests are required to determine if these glycans can distinguish lesions that will become cancerous from lesions that will remain dormant. Secondly, one of the most significant limitations of cancer biomarkers is their poor ability to differentiate tumor from benign diseases. In the examples of whole tissue blocks provided, the overlay of Hex6dHex1HexNAc5 and Hex6HexNAc2 is also able to differentiate tumor from chronic pancreatitis (Figure 39c). As CA19-9 is limited from a clinical perspective by its poor specificity in distinguishing malignant from benign conditions, this finding is particularly promising.

It is unlikely that MALDI-IMS will be used clinically for the early detection of pancreatic cancer, as tissue biopsy samples are extremely invasive and are unlikely to be collected before the disease is too advanced. Early detection disease markers for pancreatic cancer should utilize biological fluids. However, initial targeted approaches on tissue sections can reduce the complexity associated with plasma analysis (471). While no plasma from non-diseased patients was tested, glycomic profiling was unable to distinguish pooled plasma from pancreatitis and pancreatic cancer patients (Figure 41a,b). High-mannose N-glycans and members of the glycan biomarker panel were noticeably absent from plasma samples. However, these glycans were observed in the ethyl esterified N-glycans following extraction from tissue blocks, verifying that the derivatization allows for the detection of these glycoforms (Figure 41c). Thus, trends observed by tissue imaging are not reflective of the plasma glycome. However, it is unclear whether the absence of important glycoforms is attributed to their absence on plasma proteins or suppression caused by the large dynamic range of proteins in the plasma. Depletion of abundant glycoproteins from the plasma may enable the detection of low abundance glycoforms (472,473). While analysis of the plasma glycome does not correlate with our imaging results, the imaging data mirrored that reported in serum from PDAC patients, which exhibited overall elevated levels of glycosylation but no differences in high-mannose and hybrid N-glycans (385).

In conclusion, a novel disease marker discovery platform was developed that utilizes MALDI-IMS of N-linked glycans in a high-throughput TMA study. Expression levels of individual and panels of N-glycans were shown to be capable of differentiating tumor from non-tumor tissue, as well as other complex pathologies. Furthermore, the developed platform described herein can be applied to any disease for initial screening purposes. Additionally, the consistency of glycosylation trends between our tissue imaging results and published literature on serum, suggest that these glycan biomarker panels may be successful in non-invasive serum screening for pancreatic cancer.

Chapter 5: Application of Novel Glycoproteomic Approaches Directed at Pancreatic Cancer

5.1 Introduction

The ability to distinguish pancreatic cancer from non-diseased specimen by aberrations in glycosylation has been established, both herein and elsewhere. To date, the glycan antigen CA19-9, represents the only FDA approved blood test for the management of pancreatic cancer (332,333). Profiling the glycome has identified other glycoforms, including N- and O-linked glycoforms, that are differentially expressed in pancreatic cancer serum and tissue specimen (331,373,385,474,475). In tissue sections, we have recently reported experimental results from a high-throughput MALDI-IMS approach that demonstrate the potential of N-glycans to distinguish tumor from non-tumor tissue specimen with high sensitivity and specificity. These alterations in the total glycome are reflective of a pattern of altered glycosylation across many proteins (156). These glycan discriminators offer insights into the tumorigenic process, but it is doubtful that they can serve as early stage biomarkers in biological fluids due to a lack specificity for pancreatic cancer. Unlike individual glycoforms, glycoproteins that originate from a specific tissue have had clinical success as early stage cancer biomarkers. For example, alpha-fetoprotein (AFP) and prostate-specific antigen (PSA) originate from, and can detect the presence of, liver and prostate cancer, respectively (160,161,349). While profiling the glycoprotein expression level is informative, linking aberrant glycosylation to the corresponding protein can improve the clinical utility of biomarkers. The most promising example of leveraging both protein expression and glycoform occupancy in a single biomarker is AFP-L3, which outperforms AFP due to the elevation of fucosylated AFP in malignant liver diseases compared to non-malignant liver diseases (160,350,351). The identification of similar glycoproteins for pancreatic cancer to aid in the early identification of the disease would greatly improve pancreatic cancer treatment options and survival.

Early proteomic profiling of pancreatic cancer specimen revealed a general upregulation of glycoproteins in pancreatic cancer; especially those involved in regulating cellular

organization, adhesion, transport and immune system response (364,377). These observations were confirmed by global hydrazide enrichment of glycoproteins from pancreatic cancer tissue specimen, which identified 252 proteins that had aberrant expression of N-glycopeptides (283). However, both pancreatic cancer and pancreatitis exhibit similar trends in glycoprotein expression levels, with a limited number of glycoproteins that are uniquely expressed in the cancer specimen (283,380). Thus, it is important to link glycoforms to the glycoproteins in order to distinguish malignant disease from normal and benign specimen. To investigate specific glycoproteins as disease markers, most approaches utilize trends in aberrant glycosylation to determine the appropriate enrichment strategy. Efforts have been directed at linking CA19-9 to protein carriers. Mucins have long been reported to be the primary carrier of the CA19-9 antigen, but CEACAM proteins, agrin, galectin-3-binding protein, ezrin, tPA, cathepsin D, necdin, and neutrophil cytosol factor 2 are now known to carry the antigen (273,373,476). While CA19-9 is thought to be present on N- and O-glycans, it is unknown what specific glycoforms are modified with the glycan antigen (273,334,335). Other glycoproteomic approaches have assessed differential glycosylation in pancreatic cancer by selectively enriching glycan epitopes using lectin enrichment. The lectin enrichment of glycoproteins bearing specific glycan motifs, such as α2,6 sialic acids (SNA) or fucose residues (AAL), followed by downstream proteomic analysis of the enriched glycoproteome, identified glycoproteins that could differentiate pancreatic cancer from non-cancer specimen (318,343,384). Given the poor overlap in glycoprotein enrichment across various enrichment strategies, it is likely that other disease markers remain undiscovered.

Recent advances to the field of glycoproteomics can lead to a deeper understanding of the role that glycosylation plays in pancreatic cancer and the identification of disease markers. From a targeted glycoproteomics standpoint, there is a significant need to continue to identify proteins that are modified by the CA19-9 antigen. As CA19-9 is expressed in malignant and benign diseases, identification the carbohydrate antigen on proteins unique to cancer would improve the

sensitivity and specificity of disease detection. Furthermore, CA19-9 is not the only aberration in pancreatic cancer glycosylation. The global assessment of all glycoproteins and their attached glycoforms would also aid in disease marker identification. However, due to the inherent complications of glycoproteomics, these studies are rare. Intact glycoprotein analysis of complex samples requires appropriate glycopeptide enrichment steps to overcome limitations of dynamic range, reduced ionization efficiency and heterogeneity associated with glycosylation (221,264,265). Many of these studies use lectin enrichment, which limits analysis to specific glycoforms.

In this manuscript, we aim to utilize novel mass spectrometry based approaches to combine glycomic and proteomic information as a mechanism to identify disease marker candidates that can be further assessed as clinical markers for pancreatic cancer. In one approach, the incorporation of azide-labeled sialic acid analogs into pancreatic cancer cell lines provides a highly efficient enrichment strategy for sialic acid containing glycopeptides. When applied to cell lines with well characterized CA19-9 expression levels, significant variations in protein expression levels highlight a population of proteins that can be further assessed as CA19-9 carriers. In this study, Panc1 and Miapaca2 cell lines express low levels of CA19-9, while BXPC3 and Capan2 cell lines express high levels of CA19-9. As few studies have profiled the significance of N-glycans in terms of their ability to carry the CA19-9 motif, a general profile of the N-glycome is also reported in relation to the CA19-9 expression. In addition to inferences about CA19-9 expression, the glycoproteome reported within catalogs sialylated glycoproteins in pancreatic cancer. The second aim is directed at advancing methodologies available for intact glycopeptide analysis. Starting with the analysis of glycoprotein standards, we outline a workflow that can be used for the comprehensive analysis of an enriched glycoprotein or adapted for broad glycoproteomic research. As opposed to commonly employed lectin enrichment strategies, the glycopeptide enrichment approach herein is relies on physiochemical properties of the glycan,

enabling an unbiased enrichment, regardless of the glycoforms. Protocols using hydrophilic interaction chromatography (HILIC) and strong anion exchange (SAX) are outlined in the sections below.

5.2 Materials and Methods

5.2.1 Materials

Cell culture media were purchased from Mediatech, Inc. (Manassas, VA). Trypsin-EDTA used for cell culture and FBS for media supplementation were purchased from Gibco (Grand Island, NY). A mixture of Penicillin-Streptomycin-Amphotericin B was purchased from MP Biomedicals (Solon, OH). The click sugar analogs and the alkyne-agarose enrichment kit came from Life Technologies (Grand Island, NY). HPLC grade methanol, HPLC grade ethanol, HPLC grade acetonitrile, HPLC grade chloroform, HPLC grade water, histological grade xylenes, PBS, .5mL Zeba spin columns, and ammonium bicarbonate were obtained by Fisher Scientific (Pittsburgh, PA). SDS, EDTA, TrisBase, DTT, 1-hydroxybenzotriazole hydrate (HOBt), sodium hydroxide (NaOH), trypsin, calcium chloride ($CaCl_2$), fetuin, transferrin and asialofetuin were obtained from Sigma-Aldrich (St. Louis, MO). 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) was purchased from Oakwood Chemical (West Columbia, SC). NP-40 was purchased by Boehringer Manheim Gmbh (Germany). SAX solid phase extraction cartridges and C18 ZipTips came from Waters (Milford, MA) and Millipore (Ireland), respectively. Hydrazide coupling buffer was purchased by Bio-Rad (Hercules, CA) and SiMag-Hydrazide beads came from Chemicell GmbH (Berlin, Germany).

5.2.2 Cell Culture

Pancreatic cancer cell lines BXPC3, Capan2, Miapaca2, and Panc1 were provided generously by Dr. Charles Smith and were cultured as suggested by the American Type Culture Collection (ATTC). Both Panc1 and Miapaca2 cells were cultured with Dulbecco's Modified

Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and a 1X solution of Penicillin-Streptomycin-Amphotericin B. BXPC3 and Capan2 cell lines were cultured in RPMI Medium supplemented with 10% fetal bovine serum and a 1X solution of Penicillin-Streptomycin-Amphotericin B. All cell lines were maintained at 37°C in a 5% CO_2/air environment.

5.2.3 Glycomic Analysis of Cell Lines

5.2.3.1 N-Glycan Release by PNGaseF

For N-glycan analysis of the cell lysate, cells were cultured to 80% confluency and washed with PBS, prior to being scraped and pelleted. Cell pellets were lysed with 200µL of lysis buffer (1% SDS, 5mM EDTA and 25mM Tris Base) for 1hr at 4°C. The cell lysate was sonicated for 10 seconds using a probe sonicator at 30% power. Protein from 100µL of cell lysate was precipitated by the addition of 400µL methanol, 100µL chloroform, and 300µL water, followed by centrifugation at 14,000 x g for 2 minutes. After discarding the top aqueous layer, the proteins were pelleted by the addition of 400µL methanol and centrifugation at 14,000 x g for 3 minutes. The proteins were resuspended in a 1% SDS solution and reduced in 10mM DTT at 95°C for 5 minutes. After cooling to room temperature, 2.5µL NP40 and 2.6µg PNGaseF were added to the proteins from each cell line and digestion took place at 37°C overnight. Released glycans were collected by repeating the chloroform-methanol precipitation as described above and collecting the upper aqueous layer. The glycans were concentrated by vacuum centrifugation prior to beginning the ethyl esterification protocol.

5.2.3.2 Glycan Derivatization

The ethyl esterification protocol for released glycans was adapted from Reiding et al. and performed as described in Section 3.4.2.5 (237).

5.2.3.3 Data Processing

Glycan ethyl esterification of cell lines was performed in two experimental replicates for each cell line. Spectra were loaded into Compass DataAnalysis (4.0, Bruker Daltonics) and spectra were exported as .xy files for peak picking using Progenesis MALDI (Nonlinear Dynamics, Durham, NC). After the first round of peak picking, peaks that corresponded to known glycans were selected for further data processing. Other peaks corresponding to glycans that were not initially picked by Progenesis were manually added. For data visualization purposes, a heat map was generated that compares the ion intensity of a mass in a specific cell line to the average ion intensity across all cell lines. In the heat map, clades of interest were selected and individual glycan structures were identified based off of mass accuracy using GlycoWorkbench.

5.2.4 Click Chemistry Enrichment of Glycoproteins

5.2.4.1 Sugar Incorporation

Cell lines were cultured under conditions as described in section 5.2.2. For click chemistry analysis of cell lines using azide-labeled sugar analogs, cell lines were cultured to 50% confluency prior to treatment with the sugar analog. After 48 hours, the cells were washed with PBS, collected by scraping in 1mL PBS, and pelleted by centrifugation at 1,500 x g for 2 minutes. For each experimental replicate, separate plates from each cell line were treated with the 20nM ManNAz sugar analog, 20nM GalNAz sugar analog, or a corresponding volume of DMSO.

5.2.4.2 Click Chemistry Enrichment

The click chemistry-mediated glycoprotein enrichment as described by the invitrogen protocol was scaled down to adjust for a smaller number of starting cells and performed for the capture of azide-modified proteins with alkyne-agarose beads (294). Briefly, lysis buffer was added to cells, which then sat on ice for 10 minutes prior to 6 x 3 second sonications at 30% power with a probe sonicator. The lysed cells were centrifuged briefly to remove cell debris and

the cell lysate was then added to prepared alkyne-agarose resin. Glycoprotein binding occurred following the addition of the supplied catalyst solution by end-over-end rotation overnight. After overnight incubation, all non-azide modified proteins were removed by discarding the supernatant after centrifugation and washing the resin with H_2O. Glycoproteins were reduced in the provided SDS wash buffer supplemented with 10mM DTT at 70°C for 15 minutes. The wash buffer was removed following centrifugation and alkylation occurred by the addition iodoacetamide followed by incubation in the dark for 30 minutes. Bound glycoproteins were transferred to the provided spin columns and washed with SDS wash buffer, 8M Urea/100mM Tris, and 20% ACN as described in the protocol. Glycoproteins were digested with trypsin during an overnight incubation at 37°C. Released peptides were collected, reconcentrated and desalted using C18 ZipTips prior to LC-MS/MS. In total, two experimental replicates were performed and at least two technical replicates for each experimental replicate group were processed by LC-MS/MS. Following a thorough removal of nonspecific interactions, bound N-glycopeptides were released overnight during an incubation with 2.6µg PNGaseF at 37°C. Released off-bead peptides were collected and desalted using a C18 ZipTip prior to LC-MS/MS.

5.2.4.3 LC-MS/MS and Bioinformatics

A single dimension RPLC separation was employed prior to tandem MS. Dried peptides were resuspended in 10µL Buffer A (97.8% HPLC grade H_2O, 2% acetonitrile, 0.2% formic acid), injected using a Dionex U3000 nano-LC system (Thermo Scientific), and passed through a C18 trap column (C18 pepMap 100, 5µm, 300 µm x 5mm ID, Thermo). From the trap column, peptides proceeded to an in-house C18-reverse phase analytical column (.0075mm ID, 15cm) consisting of Waters YMC-ODS C18-AQ 5mm resin. Peptides were eluted using a linear gradient of increasing concentration of Buffer B (97.8% acetonitrile, 2% HPLC grade H_2O, 0.2% formic acid). For the analysis of trypsin digested glycoproteins following click chemistry enrichment, the flow rate was set to 200nL/min and ramped from 2% Buffer B to 50% Buffer B using a 180

minute linear gradient. Peptides released by PNGaseF incubation were eluted using a 240 minute linear gradient of 2-50% Buffer B at a flow rate of 200nL/min.

For the analysis of tryptic peptides and previously glycosylated peptides released from the alkyne-agarose resin with PNGaseF, the top 10 most abundant peptides were fragmented by CID in a data dependent fashion. The precursor ion scan was collected in the Orbitrap and CID was performed in the ion trap. The normalized collision energy was set to 35%, with a dynamic exclusion set for a repeat count of 2, repeat duration of 30 seconds and an exclusion duration of 100 seconds. Data were searched using Mascot (version 2.4.1, Matrix Science, London, UK) via Proteome Discover (version 1.4.0.288, Thermo Scientific). The data were searched against a UniProt Swiss Prot *Homo sapien* protein database (2014_07 release) with canonical and isoform entries. Proteins found in the common Repository of Adventitious Proteins (40,872 entries) were added to the protein database. The parameters were set to allow up to 2 missed cleavages, with a 10 ppm precursor mass tolerance and a .8 Da fragment mass tolerance. Carbamidomethyl (Cys) was defined to be a static modification, while protein N-term acetylation, oxidation (Met), and deamidation (Asn, Gln) were set as dynamic modifications. Data were also searched against an automated decoy database in Mascot. Normalized spectral counts were generated in ProteoIQ (version 2.3.08, Premier Biosoft, Palo Alto CA) and the protein readout was generated after specifying $\geq$ 2 peptides for a given protein and the protein false discover rate (FDR) <5%. Normalized spectral counts for each experimental replicate were produced by averaging the appropriate technical replicates from the sample (Sup Sheet 9). For data processing, if no spectral counts were observed, or if the average of the sugar analog treated condition was less than 5 times that of the DMSO treated condition, the spectral count value was artificially set to 1.

5.2.5 Intact Glycopeptide Analysis

5.2.5.1 Glycoprotein Digestion

Glycoprotein standards for the evaluation of glycopeptide enrichment strategies were dissolved in H_2O, reduced in 10mM DTT at 60°C for 30 minutes and alkylated in 25mM iodoacetamide in the dark for 30 minutes. The glycoprotein standards were then diluted with 100mM AMBIC, brought to a final concentration of 2mM $CaCl_2$, and digested with trypsin overnight at 37°C at a ratio of 1:100 trypsin:sample.

5.2.5.2 Glycopeptide Enrichment

For intact glycopeptide analysis, glycopeptides were enriched using strong anion exchange (SAX) SPE cartridges and cotton-HILIC tips. For enrichment with SAX SPE cartridges, the cartridge was washed and preconditioned with 1mL ACN, 1mL .1M triethylammonium acetate, 3mL H_2O, and 1mL 95% ACN 1% TFA. Digested peptides were resuspended in 400µL 50% ACN, .1% TFA and applied to the SPE cartridge along with 3mL 95% ACN, 1% TFA. After the sample was allowed to flow through the cartridge, the cartridge was washed with 3mL 95% ACN 1% TFA and glycopeptides were eluted with 500µL 50% ACN .1% TFA. Alternatively, cotton-HILIC tips were also used for the enrichment of glycopeptides. Peptides were first resuspended in 80% acetonitrile. The cotton-HILIC tips were washed 5x with 10µL H_2O and 3x with 10µL 85% ACN, followed by sample loading 40x, tip washing with 3x 10µL 85% ACN .1% TFA, and glycopeptide elution in 10µL H_2O.

For selective enrichment of sialylated glycopeptides by hydrazide enrichment, glycoproteins were incubated with 2mM periodic acid on ice for 10 minutes. Oxidation was quenched by the addition of glycerol. Proteins were desalted with a Zeba desalting column buffer exchanged with coupling buffer, prior to overnight incubation with hydrazide magnetic beads at

room temperature. Sample washing, digestion and glycopeptide release was performed as described (286).

5.2.5.3 Glycan Release and Derivatization

Enriched glycopeptides were digested with PNGaseF overnight at 37°C. Glycans were derivatized and enriched as described section 5.2.3.2.

5.2.5.4 LC-MS/MS and Bioinformatics

The same configuration of the LC-MS/MS, including the trap and analytical columns, were used as described in section 5.2.4.3. Peptides were eluted using a 120 minute linear gradient of 3-50% buffer B.

Analysis of intact glycopeptides proceeded through a product dependent ionization strategy, termed HCD-pd-ETD, on the Orbitrap Elite. The acquisition parameters were as follows. Precursor ions were detected in the Orbitrap with the resolution set to 60,000 with an m/z range from 400 - 2000. The top 10 most abundant ions were selected and fragmented by HCD and detected in the Orbitrap at a resolution of 15,000. For HCD fragmentation, the values for the minimum signal required, isolation width, normalized collision energy, and activation time were set to 500.0, 2.00, 35.0, and .100, respectively. The detection of glycan specific oxonium ions among the top 30 most abundant fragment ions triggered ETD fragmentation with supplemental activation enabled on the precursor ion. The parameters for the minimum signal required, isolation width, and normalized collision energy remained consistent between the HCD and ETD fragmentation parameters.

Byonic software version 2.2.9 (Protein Metrics, San Claros, CA) was used for intact glycopeptide identification following LC-MS/MS analysis. Briefly, data files were directly loaded into Byonic and searched against the appropriate protein database files. The parameters for the search were set as follows. Cleavage sites for trypsin digest were set to RK, with the C-terminal

cleavage side, semi-specific N-ragged digestion specificity and missed cleavages were set to -1. The precursor mass tolerance, fragment mass tolerance (HCD) and fragment mass tolerance (ETD) were set to 6ppm, 10ppm, and .6 daltons, respectively. Carbamidomethylation (Cys) was set as a fixed modification, oxidation (Met) and methylation (Cys) were set as common modifications, and pyroglutamate (pyroGlu) formation of N-terminal glutamate and glutamic acid were set as rare modifications. In total, the analysis was established to allow a maximum of two common modifications and 1 rare modification. N-glycosylation modifications were listed as a common modification in the search. For intact glycopeptide analysis, an N-glycan library was selected from Byonic or manually entered if glycomic identification had been previously performed. Data were viewed using the generated excel spreadsheet or the Byonic Viewer software. To confirm the presence of a glycopeptide, the ETD spectra and preceding HCD spectra were manually assessed in the LC-MS/MS analysis using the Xcaliber software 2.2 SP1.48 (Thermo Scientific, Pittsburgh, PA) to ensure the presence of oxonium ions in the HCD spectra.

5.3 Results

5.3.1 Glycomic Profile of Pancreatic Cancer Cell Lines

Profiling of the N-glycome was performed following ethyl esterification derivatization and HILIC enrichment. The analysis of four cell lines with characterized CA19-9 expression levels revealed general trends in glycan expression that correlated to CA19-9 status. In all cell lines, the peak corresponding to Man8 was the largest peak detected in the spectra and serves as a point of comparison (Figure 42). Importantly, the low CA19-9 cell lines exhibited a higher relative intensity of the Man8 peak and more intense high molecular weight ions ($m/z > 2750$), which correspond to branched, multiply sialylated N-glycans (Figure 42a,b). Alternatively, the high CA19-9 expressing cell lines exhibit more complexity in the MALDI spectra between $m/z = 2000$ and $m/z = 2500$ (Figure 42c,d). A molecular heat map was generated to visualize glycan

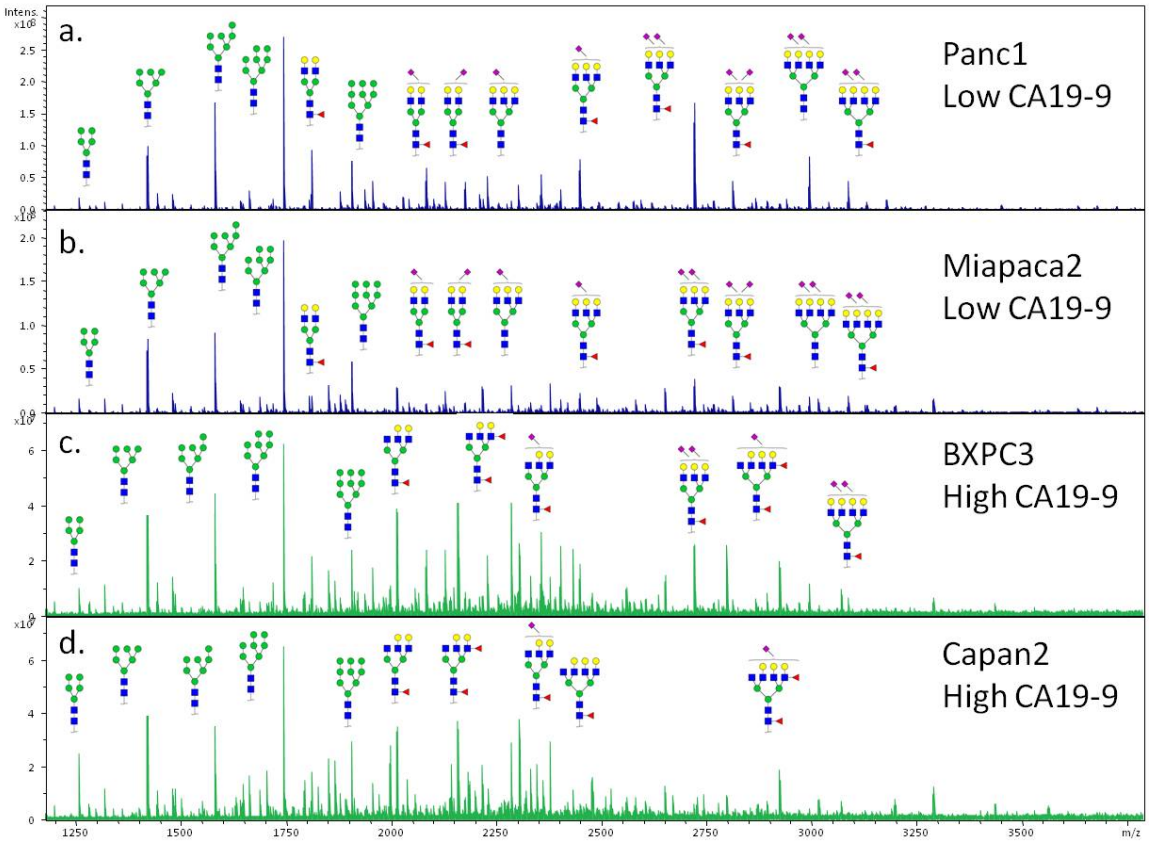


Figure 42. N-Glycan Profiling of Four Pancreatic Cancer Cell Lines. The N-glycans from four pancreatic cancer cell lines with characterized CA19-9 expression levels were released and derivatized by ethyl esterification. Noticiable differences in the annotated spectra were observed between the high CA19-9 expressing cell lines (c-d) and the low CA19-9 expressing cell lines (a-b).

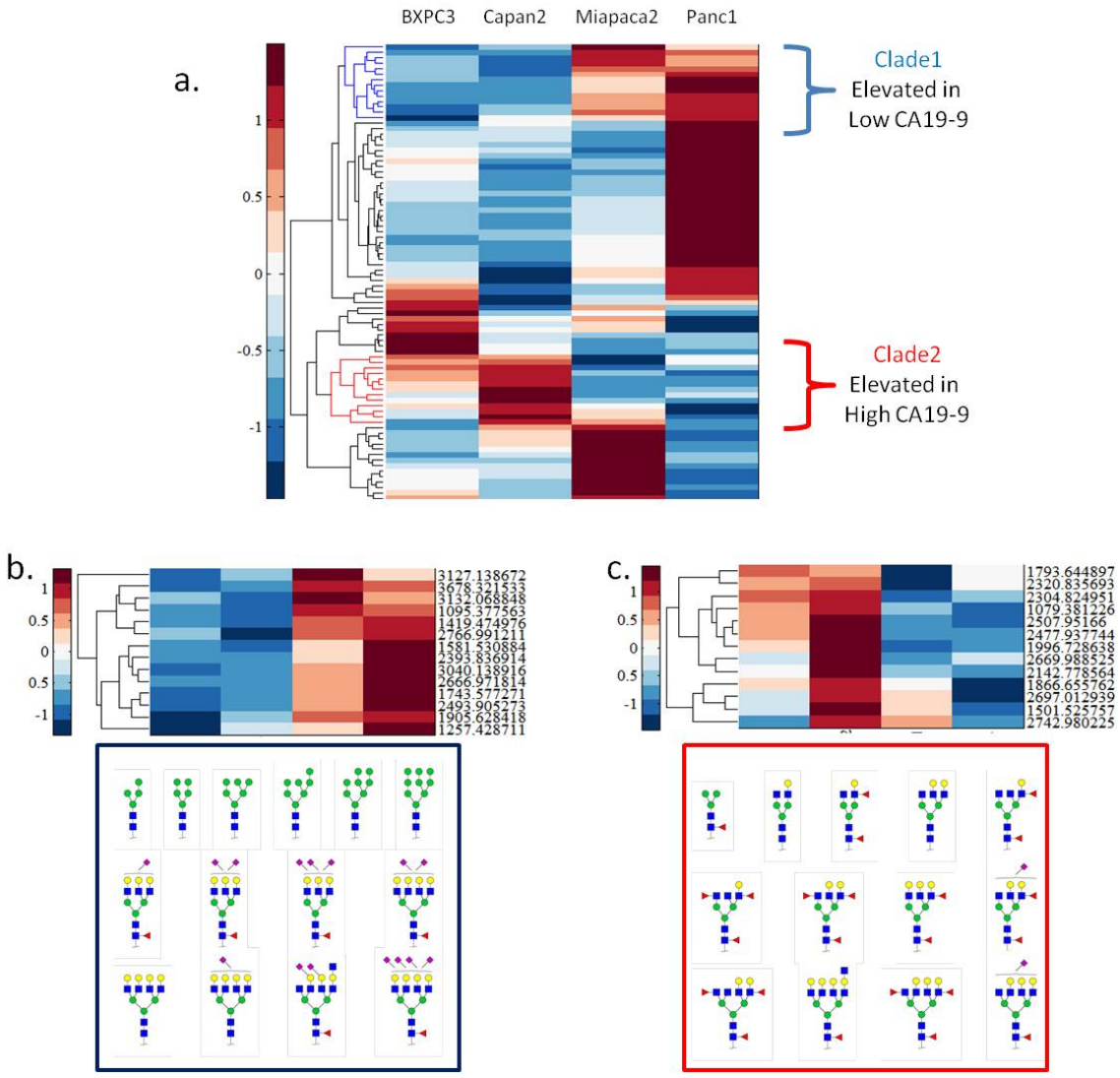


Figure 43. Molecular Heat Map of N-Glycans from the Cell Lines. Individual glycoforms were detected in progenesis and were compiled in a heatmap comparing the average intensity of an individual glycoform from one cell line, the the average across all four cell lines (a). Clades were identified of glycoforms that were elevated in both of the low CA19-9 expressing cell lines (b) or the high CA19-9 expressing cell lines (c). Glycoforms of interested were drawn using GlycoWorkbench software.

abundance in regards to CA19-9 expression levels (Figure 43). Clades that displayed consistent alterations that corresponded with CA19-9 levels were further assessed to look at the broad N-glycan structural features. In the low CA19-9 expressing cell lines, a general trend of elevated high-mannose N-glycans was observed, as well as sialylated tri- and tetraantennary N-glycans (Figure 43b). In the high CA19-9 cell lines, the most predominant trend was the presence of multiply fucosylated N-glycans (Figure 43c).

5.3.2 Sialylated Glycoprotein Profile of Pancreatic Cancer Cell Lines

Glycoprotein enrichment by click chemistry followed by proteomic analysis revealed numerous glycoproteins present in the pancreatic cancer cell lines, many of which exhibited trends that coincided with CA19-9 expression levels. In all, 535 unique proteins were identified in at least one ManNAz treated cell line. Of these 535 proteins, 429 (80.2%) were identified as glycoproteins by searching the Uniprot human proteome for the glycoprotein PTM keyword (KW-9991). Data were further processed by their group-averaged normalized spectral counts to compare the protein fold change between the high and low CA19-9 expressing cell lines. Using this approach, 102 proteins were elevated in the high CA19-9 cell lines, while 97 proteins were elevated in the low CA19-9 cell lines. The Uniprot accession IDs for both classes of proteins were assessed by WebGestalt (http://bioinfo.vanderbilt.edu/webgestalt/) using both GO and KEGG enrichment analysis to identify the top 10 most significantly enriched pathways. Among interesting findings, both GO and KEGG enrichment analysis identified pathways involved in O-glycan biosynthesis and processing to be enriched in the high CA19-9 cell lines (Table 6).

Upon manual inspection, known CA19-9 carriers were identified among the proteins that were elevated in the high CA19-9 expressing cell lines. These proteins include members of the mucin and CEACAM family of proteins, as well as galectin-3-binding protein. The data were further processed to identify glycoproteins that were exclusively detected in both high CA19-9 expressing cell lines. In all, 37 proteins met this criteria in at least one experimental replicate, of

Table 6. WebGestalt Pathway Analysis. The proteins elevated in the high CA19-9 cell lines and the low CA19-9 cell lines were processed by WebGestalt to do KEGG and GO Pathway Analysis

High CA19-9 Pathways

KEGG Pathway	ID	AdjP
Mucin type O-Glycan biosynthesis	512	2.010E-07
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	5412	6.990E-06
Complement and coagulation cascades	4610	6.990E-06
ECM-receptor interaction	4512	2.000E-04
Dilated cardiomyopathy	5414	2.000E-04
Protein digestion and absorption	4974	2.000E-04
Adherens junction	4520	2.000E-04
Hypertrophic cardiomyopathy (HCM)	5410	2.000E-04
Regulation of actin cytoskeleton	4810	3.000E-04
Vibrio cholerae infection	5110	8.000E-04

GO: Biological Process	GO	AdjP
Cell Adhesion	7155	2.280E-08
Biological Adhesion	22610	2.280E-08
O-glycan processing	16266	1.260E-05
protein O-linked glycosylation	6493	6.080E-05
homophilic cell adhesion	7156	4.000E-04
adherens junction organization	34332	7.000E-04
cell-cell adhesion	16337	7.000E-04
cell junction assembly	34329	1.000E-03
cell junction organization	34330	2.300E-03
cell-matrix adhesion	7160	2.900E-03

Low CA19-9 Pathways

KEGG Pathway	ID	AdjP
Lysosome	4142	8.610E-06
Metabolic pathways	1100	1.530E-05
Starch and sucrose metabolism	500	1.300E-03
Folate biosynthesis	790	1.300E-03
Axon guidance	4360	1.300E-03
Renin-angiotensin system	4614	3.200E-03
Small cell lung cancer	5222	3.200E-03
ECM-receptor interaction	4512	3.200E-03
Glycosaminoglycan biosynthesis - chondroitin sulfate	532	3.600E-03
Amoebiasis	5146	5.100E-03

GO: Biological Process	GO	AdjP
Cell Adhesion	7155	9.210E-09
Biological Adhesion	22610	9.210E-09
blood vessel morphogenesis	48514	1.520E-07
vasculature development	1944	2.250E-07
blood vessel development	1568	6.350E-07
cell migration	16477	1.890E-06
angiogenesis	1525	2.720E-06
localization	51179	3.540E-06
cellular component movement	6928	4.330E-06
localization of cell	51674	4.330E-06

Table 7. Exclusive Identification of Glycoproteins in High CA19-9 Expressing Cell Lines. Proteins that were uniquely identified in the high CA19-9 cell lines (at least one replicate) were compiled.

Sequence Id	Gene Symbol	Gene Name	Rep1	Rep2
P13688	CEACAM1	carcinoembryonic antigen-related cell adhesion molecule 1	yes	yes
Q9BUF7	CRB3	crumbs homolog 3	yes	yes
Q9Y653	GPR56	G protein-coupled receptor 56	yes	yes
O60449	LY75	lymphocyte antigen 75	yes	yes
Q8N271	PROM2	prominin 2	yes	yes
Q6IA17	SIGIRR	single immunoglobulin and toll-interleukin 1 receptor (TIR) domain	yes	yes
Q11201	ST3GAL1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1	yes	yes
Q9Y5Y6	ST14	suppression of tumorigenicity 14 (colon carcinoma)	yes	yes
O60637	TSPAN3	tetraspanin 3	yes	yes
P09758	TACSTD2	tumor-associated calcium signal transducer 2	yes	yes
Q96IQ7	VSIG2	V-set and immunoglobulin domain containing 2	yes	yes
Q5XXA6	ANO1	anoctamin 1, calcium activated chloride channel	no	yes
O15438	ABCC3	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	no	yes
P12830	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	no	yes
P06731	CEACAM5	carcinoembryonic antigen-related cell adhesion molecule 5	no	yes
P40199	CEACAM6	carcinoembryonic antigen-related cell adhesion molecule 6	no	yes
O15551	CLDN3	claudin 3	no	yes
Q96PD2	DCBLD2	discoidin, CUB and LCCL domain containing 2	no	yes
Q9BS26	ERP44	endoplasmic reticulum protein 44	no	yes
P98172	EFNB1	ephrin-B1	yes	no
P16422	EPCAM	epithelial cell adhesion molecule	no	yes
P31943	HNRNPH1	heterogeneous nuclear ribonucleoprotein H1 (H)	yes	no
P56199	ITGA1	integrin, alpha 1	no	yes
P80188	LCN2	lipocalin 2	no	yes
Q86X29	LSR	lipolysis stimulated lipoprotein receptor	yes	no
Q04912	MST1R	macrophage stimulating 1 receptor	no	yes
Q5SRI9	MANEA	mannosidase, endo-alpha	no	yes
Q15223	PVRL1	poliovirus receptor-related 1	no	yes
Q16651	PRSS8	protease, serine, 8	yes	no
Q8TCT8	SPPL2A	signal peptide peptidase like 2A	no	yes
Q8WWI5	SLC44A1	solute carrier family 44, member 1	no	yes
Q92485	SMPDL3B	sphingomyelin phosphodiesterase, acid-like 3B	no	yes
Q9NRS4	TMPRSS4	transmembrane protease, serine 4	no	yes
P62987	UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1	yes	no
Q14435	GALNT3	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3 (GalNAc-T3)	yes	no
Q14508	WFDC2	WAP four-disulfide core domain 2	yes	no

which 11 proteins satisfied these conditions in both experimental replicates (Table 7), including CEACAM1, CEACAM5, and CEACAM6. To assess the contribution of N-linked glycosylation to the glycoprotein enrichment, bound N-glycopeptides were released with PNGaseF and analyzed by LC-MS/MS (Sup Sheet 8). A majority of the proteins were not identified after PNGaseF release, emphasizing the importance of O-linked glycosylation. Interestingly, galectin-3-binding protein, was among the proteins identified after PNGaseF.

5.3.3 Optimization of Intact Glycopeptide Analysis

Fetuin, asialofetuin and transferrin glycoprotein standards were utilized to validate the experimental and instrumental parameters for intact glycopeptide identification by LC-MS/MS. Initially, to confirm the enrichment of glycopeptides over non-glycosylated peptides from the glycoprotein standard, a combination of LC-MS/MS and MALDI-MS was used. For this process, peptides were either desalted using a C18 ZipTips or glycopeptides were enriched by SAX SPE cartridges. Enriched glycopeptides were incubated with PNGaseF to generate previously glycosylated peptides. Peptides (Figure 44a), enriched glycopeptides (Figure 44c), and previously glycosylated peptides (Figure 44b) were all spotted on a MALDI plate with CHCA matrix. LC-MS/MS of the total peptide solution was also performed and searched against a protein database containing fetuin and transferrin. MALDI analysis of the peptide sequences revealed numerous peptides for both fetuin and transferrin that corresponded to the peptide m/z values identified by LC-MS/MS. In contrast, the ions detected following SAX enrichment displayed generally larger m/z values that did not correspond to the LC-MS/MS analysis, but likely correspond to in-source decay of the glycan moieties on the glycopeptides. This hypothesis was verified by MALDI analysis of deglycosylated peptides, which identified all five peptides known to be N-glycosylated on fetuin and transferrin. A similar trend was observed when using cotton-HILIC tips for glycopeptide enrichment as opposed to SAX SPE cartridges (data not shown). Additionally, the N-glycans were released by PNGaseF following glycopeptide enrichment to

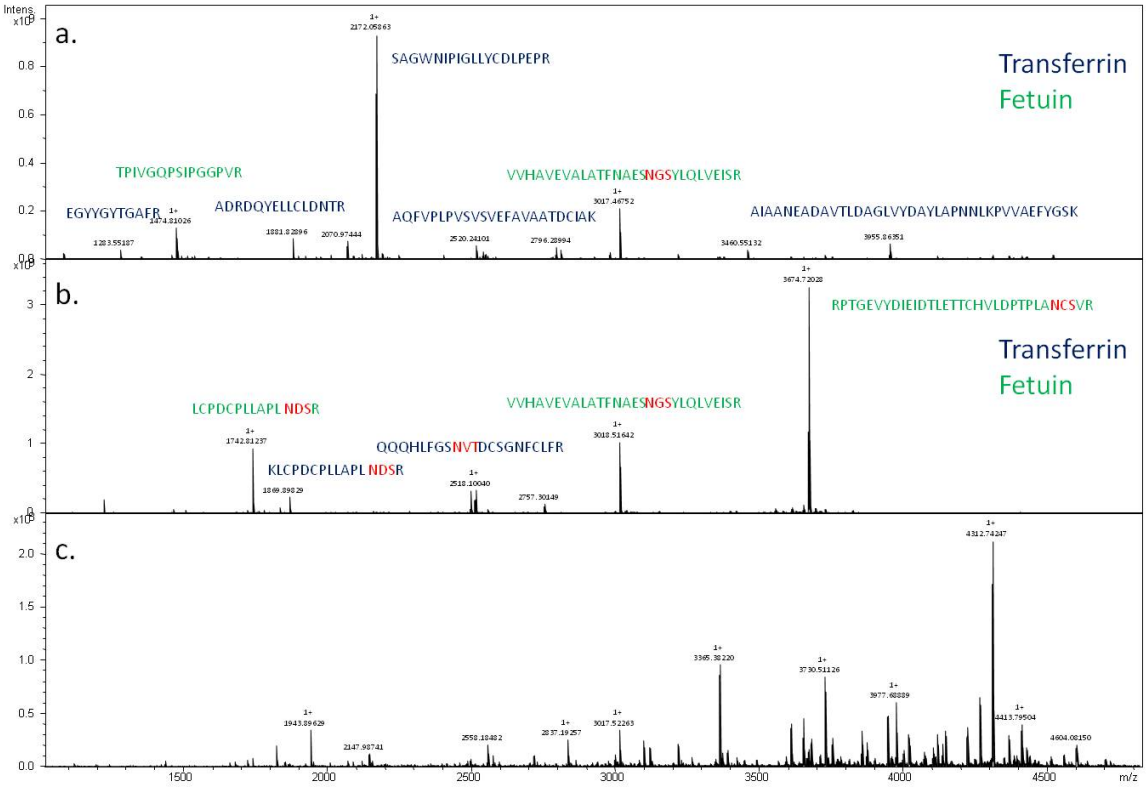


Figure 44. MALDI Analysis of Enriched Glycopeptides. To test for the efficacy of glycopeptide enrichment, peptides were enriched using C18 (a) or SAX (c) SPE cartridges. Peptides characterized from fetuin (green text) and transferrin (blue text) are annotated in the spectra. SAX enriched glycopeptides were digested with PNGaseF to release N-linked glycans, which resulted in the identification of deglycosylated peptides from all of the known N-glycopeptide sites (b). Furthermore, sites of deamidation resulting from PNGaseF release are presented in red text.

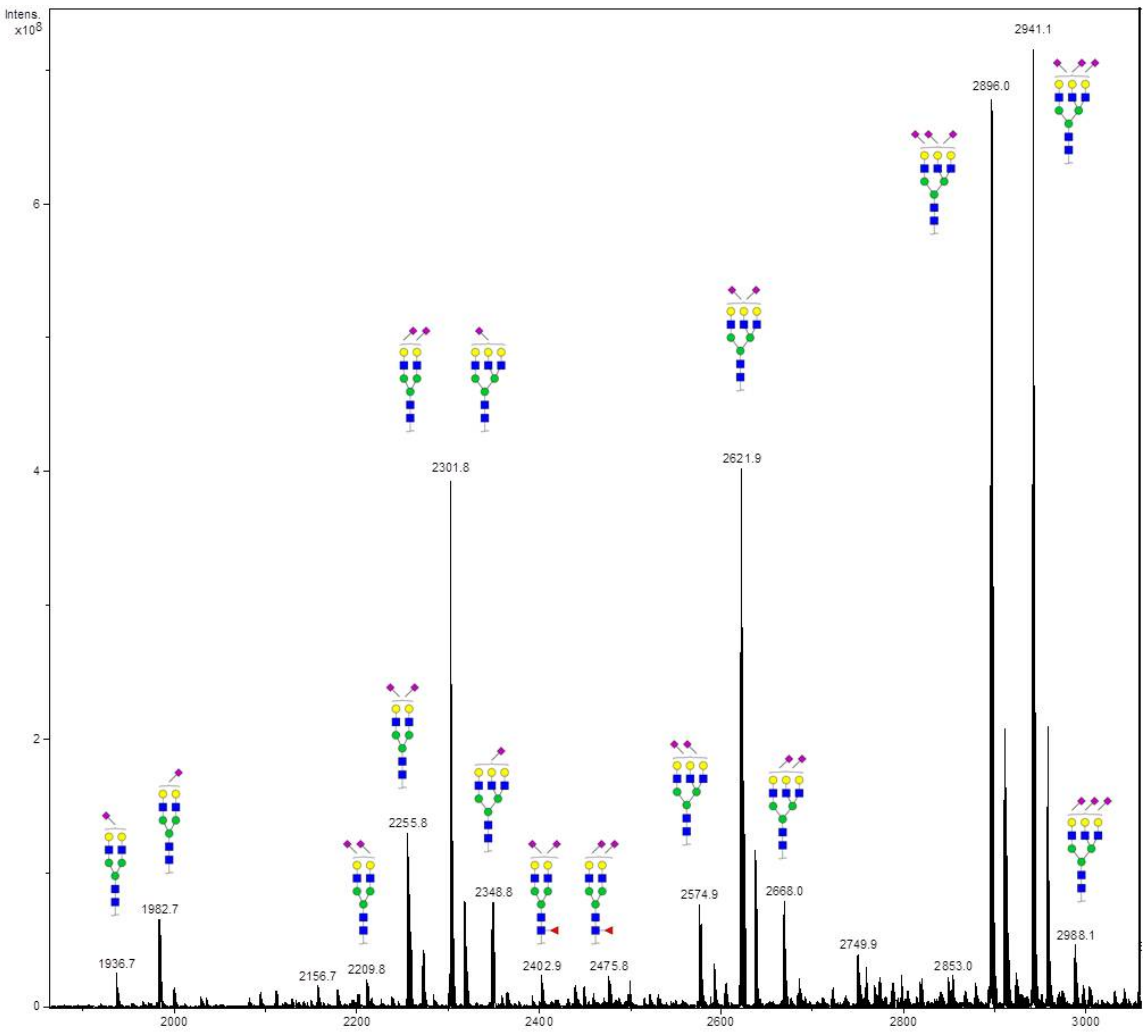


Figure 45. Glycomic Analysis of Enriched Glycopepdies from Fetuin and Transferrin. Glycans were released from enriched peptides by PNGaseF and then characterized by MALDI analysis of the ethyl esterified glycoforms. Presented is an annotated spectra of the derivatized glycoforms.

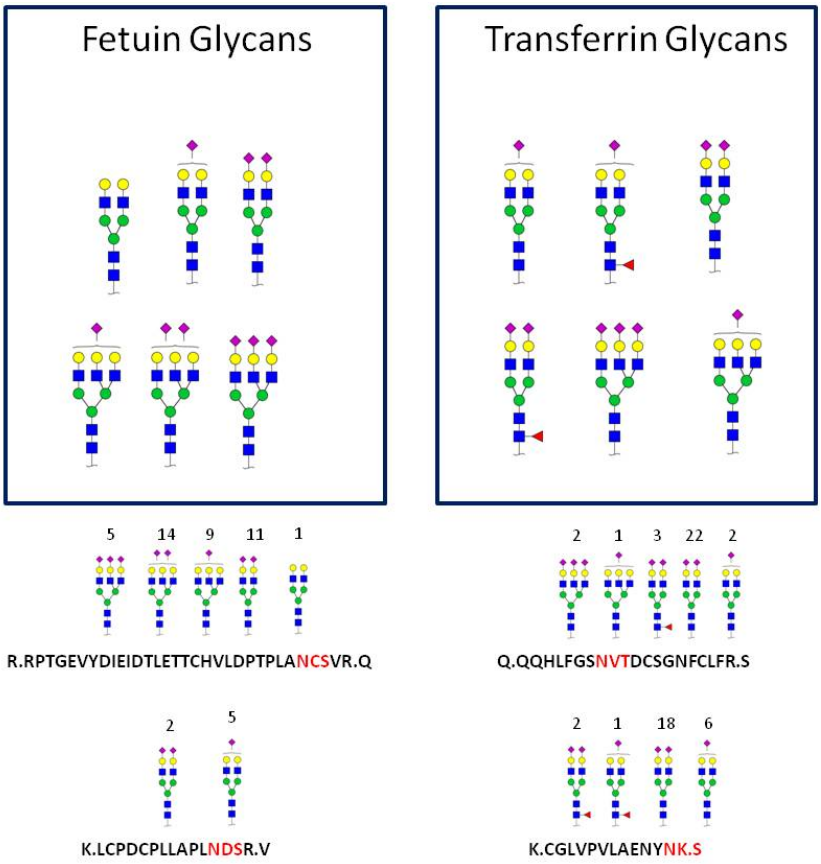
profile the corresponding glycoforms. A spectra of ethyl esterified glycans for fetuin and transferrin is provided (Figure 45). Analysis of identified glycoforms revealed primarily complex, sialylated N-glycans with some fucosylation present.

SAX enriched glycopeptides from the glycoprotein standards were then analyzed on the Orbitrap Elite using the HCD-pd-ETD ionization settings for the identification of intact glycopeptides and the corresponding glycoforms. ETD spectra were searched in Byonic to identify N-glycosylated peptides. In the Byonic search, peptides were searched against a manually defined N-glycan library consisting of identified glycans (Figure 46). In the search readout, all five sites of N-glycosylation were detected with N-glycan modifications present (Figure 47). A further interrogation of the glycopeptides after removal of all glycopeptides with a score < 50, identified six glycoforms present on both Fetuin and Transferrin. The number of times a specific glycoform was present on a specific peptide was recorded, and demonstrates N-glycoform heterogeneity (Figure 46).

In an alternative approach, glycoprotein standards were enriched using magnetic hydrazide resin. Under low oxidizing conditions, only sialic acids are oxidized and form covalent bonds with the hydrazide resin. After tryptic digest, intact glycopeptides can be released by acid hydrolysis, which cleaves the glycosidic bond between the sialic acid and terminal galactose residues. Intact glycopeptide analysis on the Orbitrap Elite identified similar glycopeptides as after SAX glycopeptide enrichment, with the exception that no sialylated glycans were identified (Figure 48).

5.4 Discussion

The carbohydrate antigen CA19-9 is the most well characterized disease marker for pancreatic cancer, but knowledge of the glycoform and glycoprotein carriers of the antigen is still incomplete. To our knowledge, no examples of specific N-glycoforms that carry the motif are known. Additionally the contribution that N-glycosylation makes to antigen expression is poorly



Glycan Composition	Glycan EE	Byonic
Hex5HexNAc4	Yes	Yes
Hex5dHex1HexNAc4	Yes	No
Hex5HexNAc4NeuAc1	Yes	Yes
Hex5dHex1HexNAc4NeuAc1	Yes	Yes
Hex5HexNAc4NeuAc2	Yes	Yes
Hex5dHex1HexNAc4NeuAc2	Yes	Yes
Hex6HexNAc5	Yes	No
Hex6HexNAc5NeuAc1	Yes	Yes
Hex6dHex1HexNAc5NeuAc1	Yes	No
Hex6HexNAc5NeuAc2	Yes	Yes
Hex6HexNAc5NeuAc3	Yes	Yes
Hex7HexNAc6NeuAc1	Yes	No
Hex7HexNAc6NeuAc2	Yes	No
Hex7HexNAc6NeuAc3	Yes	No
Hex7HexNAc6NeuAc4	Yes	No

Figure 46. N-Glycan Identification from Glycoprotein Standards. Glycans identified from the Byonic search for both fetuin and transferrin are listed, as well as the number of times each glycoform was observed on a specific peptide. The table represents a list of glycoforms that were identified in the glycomic analysis and were used in the Byonic search.

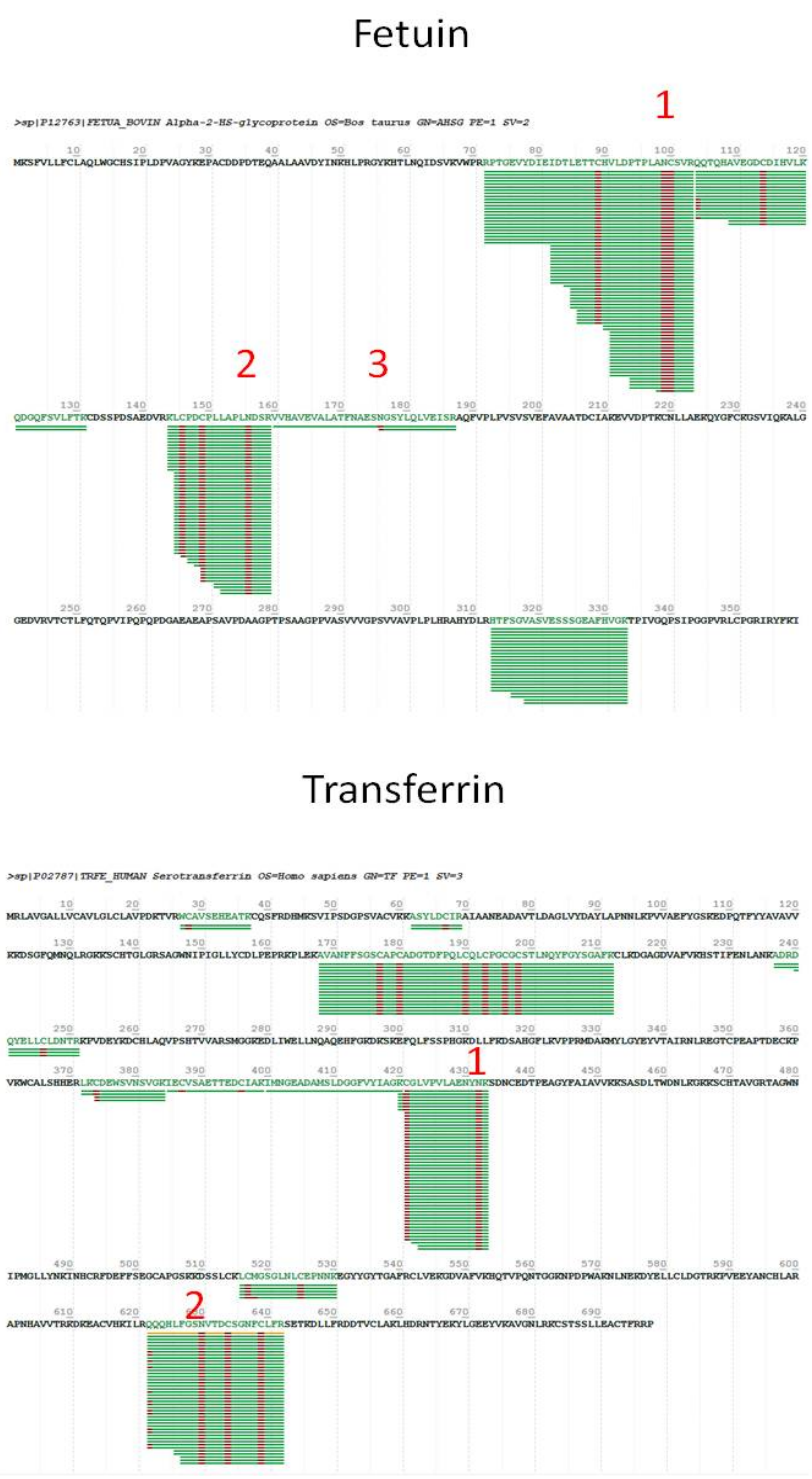


Figure 47. Peptide Sequences Identified by ETD Ionization. ETD spectra from the HCD-pd-ETD analysis of the glycoproteins were searched with Byonic. Peptides, including intact glycopeptides, from both fetuin and transferrin were identified.

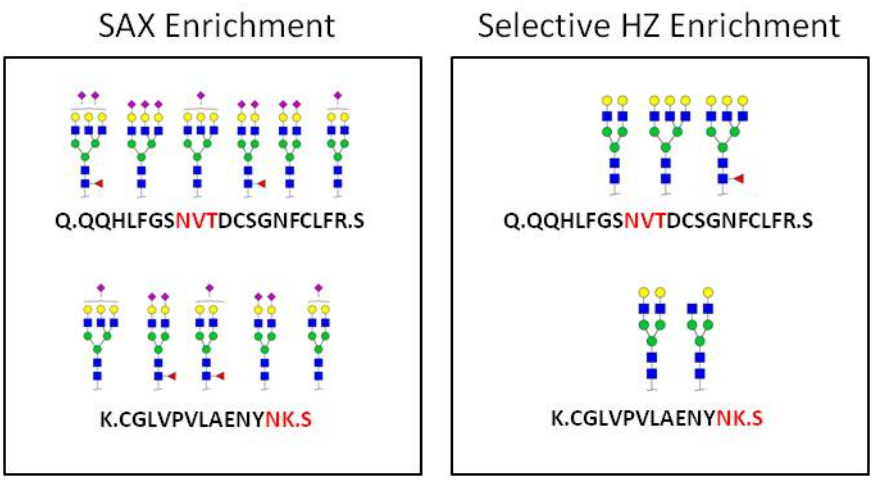


Figure 48. Comparison of N-Glycan Identification from SAX and Selective Hydrazide Enrichment. Glycoforms identified on each peptide during intact glycopeptide analysis on SAX and selective hydrazide enrichment reveal similarities in the common core structure of the glycoform. The predominant difference is that the glycoforms identified following selective HZ enrichment is that no sialylated glycoforms are observed, which is expected as glycoforms were released by acid hydrosis to cleave the sialic acids.

characterized. To address these limitations, we performed N-glycomic and glycoproteomic profiling of pancreatic cancer cell lines with well characterized CA19-9 expression levels. In the glycomic analysis, a selective derivatization strategy was used to distinguish α2,3- and α2,6-linked sialic acids by induced changes in the molecular weight. Across the cell lines, general trends in the relative abundance of N-glycans were observed between high and low CA19-9 expressing cell lines (Figure 42). Interestingly, the overall level of sialylated N-glycans was not increased in high CA19-9 expressing cell lines, as originally hypothesized. In fact, the most noticeable difference was that multiply fucosylated N-glycans were increased in high CA19-9 expressing cell lines, while high-mannose glycans were elevated in low CA19-9 expressing cell lines (Figure 43). This finding suggests that the fucosyltransferase responsible for CA19-9 expression on O-glycans also plays a role in N-glycosylation. In our data, only two of the N-glycoforms elevated in the high CA19-9 cluster contained a sialic acid, none of which were α2,3-linked. Additionally, in the glycoproteomic analysis, there were few protein identifications following PNGaseF incubation, suggesting CA19-9 is primarily present on O-glycans. Interestingly, galectin-3-binding protein and CEACAM proteins were both identified following PNGaseF release. As galectin-3-binding protein is known to express CA19-9 on N-glycoforms, further assessments of CEACAM N-glycans should be performed (273). Given the glycomic and glycoproteomic results, it is evident that CA19-9 is predominantly carried on O-linked glycans. One of the primary limitations of CA19-9 as a clinical disease marker is that a large population of individuals do not express fucosyltransferase 3 and are therefore incapable of expressing CA19-9 (333,343). Profiling the N-glycome of patients with pancreatic cancer that do not express CA19-9 should reveal a similar decrease in polyfucosylated N-glycans.

Glycoproteomic analysis of the cell lines further implicated processes related to O-glycosylation as a primary distinguisher between high and low CA19-9 expressing cell lines (Table 6). Additionally, numerous proteins that are known CA19-9 carriers were identified to be

elevated in the high CA19-9 expressing cell lines. Among these proteins, some were uniquely expressed in the high CA19-9 expressing cell lines (Table 7). Given the propensity for the technique to identify known CA19-9 carriers in the high CA19-9 cell lines, identified proteins should be tested for CA19-9 expression (273,373,476). Importantly, determining CA19-9 protein carriers has improved the performance of the antigen as a disease marker (477). Yeu et al. used a sandwich-ELISA assay to profile CA19-9 expression on selected proteins from serum samples, and found elevations in the glycan antigen on MUC5ac, even in cases where the protein level did not change (373). A receiver operating characteristic (ROC) curve showed that using CA19-9 on MUC5ac instead of MUC5ac protein expression improved the AUC from .72 to .90 (373). The same group found that combining information related to total CA19-9 expression with the level of CA19-9 on MUC5ac and MUC16 was able to improve the performance of the disease marker (478).

While inferences can be made about proteins with regard to CA19-9 expression, this data largely details and catalogs variances in glycoprotein sialylation and not CA19-9 specifically. Thus, it is likely that many of the proteins identified do not express CA19-9. Various integrins were observed to be elevated in the high CA19-9 cells and have been reported in literature as exhibiting higher levels of sialylation in pancreatic cancer. Given the dynamic role of sialic acid in tumor progression, the comprehensive library of sialylated glycoproteins identified in the click chemistry approach can be assessed in serum samples (362,363).

While click chemistry represents a promising technique for initial disease marker discovery, the primary limitation of the approach is that it does not enable analysis of the glycan motifs on specific proteins. As observed with AFP-L3, this information can be critical to improving the clinical utility of the disease marker. To adjust for these limitations, intact glycoproteomic workflows were optimized on glycoprotein standards. SAX enrichment of a mixture of fetuin, asialofetuin and transferrin was able to specifically enrich glycopeptides from

the non-glycosylated peptides, as assessed by both MALDI and LC-MS/MS (Figure 44). However, asialylated glycoforms were largely missing during glycomic and glycoproteomic analysis, suggesting SAX is an efficient enrichment strategy for sialylated glycopeptides. There was a high degree of overlap between glycans that were detected in the glycomic analysis compared to glycans identified in the glycoproteomics workflow (Figure 45). When applied to complex protein mixtures (i.e. serum, tissue, cell lysates), SAX and HILIC glycopeptide enrichment were insufficient to enrich glycopeptides from the total glycopeptide mixture present. Traditionally, glycoproteomic analysis has been limited due to issues of broad dynamic range of proteins and a limited number of glycopeptides among total peptides (221,262–264). To compensate for these limitations, glycoconjugate enrichment is required to selectively enrich glycopeptides prior to analysis by LC-MS/MS. Glycopeptide enrichment strategies that rely on physiochemical properties generally exhibit poor specificity for glycopeptide enrichment, so other approaches should be considered (310). To increase the specificity of glycopeptide enrichment, a new technique involving hydrazide chemistry under low oxidizing conditions was used to enrich sialylated glycoproteins (276,286). With this technique, intact glycopeptides can be released and surveyed, at the expense of the terminal sialic acid. In our data, the underlying glycoforms detected in the selective hydrazide enrichment and SAX glycopeptide enrichment were largely similar (Figure 48), with the main difference being the lack of sialic acid due to acid hydrolysis release. Glycoprotein enrichment using covalent bond formation to solid phase support is more efficient and could mitigate issues of protein dynamic range in the analysis of complex samples. However, this approach is only useful in characterizing sialylated glycopeptides, and not for general analysis. The intact glycoproteomic approach outlined can be further implemented to the study of pancreatic cancer. Individual glycoproteins observed in the click chemistry approach can be enriched by immunoprecipitation and profiled by LC-MS/MS to characterize the presence and occupancy of individual glycoforms. Alternatively, aberrantly expressed glycoforms and the

corresponding glycoprotein identified by MALDI-IMS could be enriched by lectin enrichment in order to identify the protein carriers.

In conclusion, this aim expanded the field of pancreatic cancer glycoproteomics by defining a class-specific role of glycans as CA19-9 expressors, identifying putative CA19-9 carrier proteins, documenting sialylated glycoproteins in pancreatic cancer, and developing methodologies for intact glycopeptide analysis of identified proteins. While intact glycopeptide analysis was performed on glycoprotein standards, future work can apply the outlined methods to isolated or total glycoproteins from biological samples. As CEACAM proteins were identified in the elevated high CA19-9 cell lines, and sialylation was detected at sites of N-glycosylation, enrichment and intact glycoprotein analysis of CEACAM proteins is a logical next step.

Chapter 6: Conclusions, Limitations and Future Studies

6.1 MALDI-IMS Optimization for N-Glycan Analysis

6.1.1 Conclusions

We hypothesized that PNGaseF application on tissue sections would release N-linked glycans and enable glycan localization, identification and relative abundance to be evaluated in a single experiment. To test this hypothesis, tissue sections from mouse kidneys and brains were used for initial optimization before further analysis of biologically relevant diseased tissues. To ensure released ions were N-glycans, a thorough characterization using up-to-date glycomic methodologies was performed. The application of these approaches revealed:

1. MALDI-IMS sample preparation techniques employed for MALDI-IMS of tryptic peptides, including washes to delipidate tissue sections and the application of a uniform layer of enzyme, can be adapted for the imaging of N-glycans directly from tissue sections.

2. The common core structure of N-linked glycans, the limited number of monosaccharides available and defined biosynthetic pathways make the characterization of the glycan composition relatively simple. Additionally, comparing ions to a serial tissue section that is shielded from enzyme application results in a robust difference in ion population, with N-glycans present only in the digested tissue.

3. N-glycan localization is often reflective of tissue histology. A majority of glycans are not observed to be homogenously distributed across entire tissue sections but correlate with specific cellular populations. N-glycans are elevated in multiple regions of a tissue, suggesting that there is a concrete difference in glycosylation that is linked to histology, as opposed to alterations being an artifact of the technique.

4. Cancer sections of complex histology reveal patterns of N-glycans that correlate to histopathology analysis. In addition to differentiating tumor from non-tumor tissue

sections, glycan signatures correlate with inflammation, fibrosis, and regions of tissue histology.

5. The composition of released N-glycans detected by MALDI-IMS can be confirmed by glycan standards, released glycans from glycoproteins, N-glycan structural libraries, HPLC analysis of 2AB-labeled N-glycans, permethylation, ethyl esterification derivatization, and CID. A comparison of glycans identified by off-tissue derivatization experiments to those from MALDI-IMS identified similar glycan compositions.

6. MALDI-IMS can effectively profile N-glycan distribution in FFPE tissue sections. The workflow for glycan analysis in these sections results in an increase in ion intensity over fresh-frozen tissue sections. Superior ion intensity may be a result of more stringent washes during the antigen retrieval workflow, which possibly reduce ion suppression attributed to lipid species. Similarly, slide mounted fresh-frozen tissue sections can be fixed in formalin and subjected to antigen retrieval to produce superior imaging results.

7. N-glycans released from fresh-frozen and FFPE tissue sections are consistent with glycomic analysis using traditional approaches. The analysis of FFPE mouse kidney sections identified all 28 N-glycans reported to be present in the Consortium for Functional Glycomics N-glycan database of the mouse kidney. Hex7HexNAc6, which we reported to be elevated in HCC tumor, has been previously reported to be elevated using IHC staining techniques.

8. Experimental throughput is greatly improved by performing experiments in TMAs. High-throughput N-glycan analysis by MALDI-IMS can identify N-glycans that are capable of distinguishing tumor from non-tumor tissue cores. In an analysis of an HCC TMA, Hex7HexNAc6 (elevated in tumor) and Hex8HexNAc2 (decreased in tumor) both exhibited an AuROC ≥ 0.8 and a p-value $< .05$.

9. There is a common trend of elevated expression of the glycome across multiple cancers cancer. In HCC, 21 out of 26 identified glycans were elevated in the tumor tissue, 13 of which were elevated at a significant level (p-value < .05). In pancreatic cancer, 20 out of 23 identified N-glycans were elevated in the tumor tissue, 18 of which were elevated at a significant level (p-value < .05). In both examples, high-mannose N-glycans were the primary glycoforms that were not elevated in the tumor tissue cores.

10. Statistical findings in TMAs are representative of N-glycan profiling of whole tissue blocks.

11. When applied to TMAs containing molecular subtypes of cancers from a single organ, glycoform variation by MALDI-IMS can distinguish the molecular subtypes. Individual N-glycans are expressed with varying abundance across different tumor subtypes.

12. Patterns of N-glycan distribution are observed in similar histologies across different organs. For example, complex biantennary glycans with various degrees of fucosylation and sialylation, are elevated in fibroconnective tissue in pancreas, renal and liver tissue.

13. Aberrations of structurally similar N-glycans can be linked to genomic analysis and glycosyltransferase expression.

14. MALDI-IMS has successfully been applied to a broad range of tissue sections: In the dissertation, examples were provided for mouse kidney, mouse brain, prostate, pancreas, renal, and liver tissues. In addition to these tissues, the method has also been applied to placenta, lung, colon, breast tissue and others.

6.1.2 Limitations and Future Research

Aim 1 limitations are grouped into four main categories; [1] the intrinsic need to optimize novel methodologies, [2] identifying detected glycoforms, [3] well characterized limitations of

MALDI-MS in glycan analysis, and [4] limitations in MALDI-IMS.

While the pre-imaging workflow has successfully released N-glycans that are amenable to imaging, sample preparation conditions will need to be further optimized. These conditions include the washing of the slide, the antigen retrieval process and buffer, the amount of PNGaseF required, and the time needed for enzyme digestion. It is likely that these conditions will vary between tissue types.

Due to the branched nature of N-glycans, the glycomic approach outlined yields the composition of the N-glycan but fails to identify glycan linkages. Distinguishing linkages between different glycoforms with the same composition is essential to understanding how the glycan functions biologically, as subtle differences can have dramatic implications for interactions with glycan-binding proteins. In order to obtain this information by mass spectrometry, N-glycans must be fragmented to generate both cross-ring cleavages and cleavages across the glycosidic bonds (247). However, MALDI ionization primarily produces cleavages across glycosidic bonds (218). Additionally, fragmentation of ions during MALDI-IMS of tissue sections is generally less effective than analysis of enriched material. A variety of approaches can be adapted in order to generate information indicative of anomeric linkages. In one approach, N-glycan imaging and N-glycan extraction, could be performed on serial tissue sections. This general approach was outlined for both permethylation and ethyl esterification in order to validate the underlying glycan composition. Derivatization, chromatographic separation and tandem MS/MS experiments on extracted glycans could generate cross-ring cleavages that could be used to make inferences about the glycan structure (479). Alternatively, sequential on-tissue exoglycosidase digestion could be used to assess linkages of ions corresponding to specific glycoforms. In initial experiments, application of sialidase on tissue in conjunction with PNGaseF cleaved terminal sialic acids and altered the glycan profile observed. Derivatization strategies that

selectively distinguish linkages of specific monosaccharides, such as ethyl esterification of sialic acids, could also be adapted to on tissue experiments.

MALDI-MS is one of the most established methods for N-glycan analysis, but is plagued by well-documented limitations. In positive ion mode, neutral N-glycans are primarily observed as the $[m + Na]^+$ adduct, but acidic glycans do not ionize efficiently suffer from dissociation of sialic acids (250). The dissociation of sialic acid residues results in overlapping peaks that artificially distort the abundance of the non-sialylated glycan. The ion source for the SolariX FTICR is under intermediate pressure, which reduces the loss of sialic acid to some extent (241). To compensate for these factors, glycomic analysis by MALDI-MS often relies on derivatized N-glycans, which functions to stabilize the labile monosaccharides and improve detection sensitivity. However, these methods often require glycan enrichment, which functions to remove excess salts and derivatization reagents, but is thus incompatible with MALDI-IMS (230). Alternative derivatization strategies that could stabilize labile sialic acid modifications while the glycan is still attached to the protein are promising alternatives. Both ethyl esterification and p-toluidine derivatization of sialic acids could be performed on tissue sections, and labeling reagents could be removed by washing prior to enzymatic digestion (235,237). As we have specifically observed in-source decay of sialic acids, selective labeling of these monosaccharides would greatly improve imaging analysis. MALDI-IMS in negative ion mode could also be used to profile sialylated N-glycoforms (480). An added benefit of glycan profiling in negative ion mode is that more cross-ring cleavages could occur by CID, yielding more information on anomeric linkages.

Given the sheer size of high resolution FTMS data, statistical approaches such as ClinPro Tools (Bruker Daltonics) do not work for biomarker identification. For this reason, an in-house script for peak picking and data mining was generated. Data processing using the in-house script has resulted in ions that do not correspond to N-glycans, but seem to be able to distinguish

between tumor and non-tumor tissues. This finding may be attributed to the effects of tissue thickness or cellular morphology differences that affect ionization but are not indicative of ion concentration. Another limitation of the peak picking process is that it uses binned windows, which are not ideal for selecting the center of the picked peak and cause artificially high PPM errors (481). Furthermore, the comparison of results from multiple imaging slides requires extreme consistency in the mass accuracy of the identified ions. Mass variation should be minimal across different time points, including across entire imaging experiments and separate imaging experiments. Eliminating mass defects will allow the peak picking approach to select identical peaks across experiments, providing more consistent data. In addition to the built in lock mass function, other studies have optimized calibration and peak picking approaches for the FTMS to allow for identification of analytes with sub-PPM mass accuracy (481).

MALDI-IMS is associated with relative concentrations, but ideal biomarker discovery techniques should utilize a more quantitative workflow. For MALDI-IMS, quantitative data requires an efficient matrix/enzyme application to generate a uniform and homogenous spray, and relies on the assumption that tissue morphology does not affect ionization efficiency. Efforts directed at identifying appropriate internal standards for normalization within and across slides are being investigated (460). To deal with the limitation of MALDI-IMS for comparing imaging analysis on different slides and working towards a quantitative analysis, a labeled or non-human N-glycan could be applied homogenously on the tissue section during the matrix application step. In a study by Pirman et al., MALDI-IMS following the incorporation of an deuterated internal standard resulted in quantitative results that were comparable to LC-MS/MS analysis (461). Alternatively a stable isotope labeled N-glycan could be applied homogenously on a tissue section to be used as an internal calibrant and normalization factor. Tao et al. developed an approach to obtain relative quantitation of N-glycans by incorporating ^{18}O-water onto the reducing end of glycans and profiling by direct infusion (482). In an alternative approach, Walker

et al. demonstrated the utility of using xylosylated plant glycoproteins to control for enzymatic digestion efficiency (483). Released glycans from the glycoproteins could be used to monitor digestion efficiency, serve as an internal calibrant, and be used for normalization across imaging experiments.

MALDI-IMS of fresh-frozen, FFPE, and formalin fixed tissue sections should be compared to assess any variations that depend on tissue specimen acquisition and storage. While, we observe higher signal in FFPE tissue sections compared to fresh-frozen tissue sections, individual trends of glycan abundance and localization should be compared. Few studies have investigated the impact the FFPE fixation has on PTMs, or more specifically glycosylation. Proteomic analysis of FFPE tissue sections is impacted by factors relating to sample storage and antigen retrieval (484). The affect that these analytical factors have on glycosylation should be determined. If processes related to tissue acquisition and initial storage conditions significantly alter glycomic profiling, special care should be taken to insure studies use tissue collected using well-defined tissue preservation protocols.

6.2 Glycomic Profiling of Pancreatic Cancer Tissue Specimen

6.2.1 Conclusions

Preliminary findings from pancreatic cancer tissue blocks revealed N-glycan patterns that correspond to tissue histology. Based off of these findings, we hypothesized that N-glycan profiling by MALDI-IMS could differentiate tumor from non-tumor tissue sections in a large high-throughput study. For this study, six TMAs consisting of patient-matched tumor and non-tumor tissue from 76 patients were surveyed. Data were then processed manually and by machine learning approaches to identify panels of N-glycans that are capable of distinguishing tumor from non-tumor tissue. The expression of selected N-glycans was validated on whole pancreatic cancer

blocks and linked to complex histopathology. By performing research relevant to this aim, we found:

1. Upwards of 90 ions corresponding to N-linked glycans can be identified in a single pancreas tissue block.

2. Individual N-glycan distribution can be linked to regions of pancreas histology. Not only can the profile of an individual glycan distinguish tumor from non-tumor tissue, but also GI tissue, fibroadipose connective tissue, smooth muscle, and more.

3. An in-house peak picking algorithm and data processing tool identified 54 monoisotopic peaks, 23 of which correspond to N-glycans, that are present across all of the TMAs. Among the identified N-glycans, nearly all are elevated in tumor compared to non-tumor tissue.

4. While individual N-glycans are able to differentiate tumor from non-tumor tissue cores, the use of a single N-glycan discriminator results in many false identifications. In contrast, the combination of multiple N-glycans in a biomarker improves the sensitivity and specificity of the assay

5. When assessing glycosylation patterns between tumor and non-tumor tissue cores, complex fucosylated N-glycans are the most elevated glycoforms in tumor tissue sections, while high-mannose N-glycans are present in similar amounts between tumor and non-tumor tissue cores. This trend is reflective of published literature performed on pancreas serum specimen.

6. A glycan biomarker panel consisting of Hex7dHex1HexNAc7 and Hex6dHex2HexNAc5 can able to outperform CA19-9 staining in terms of sensitivity and specificity for distinguishing tumor from non-tumor tissue cores.

7. In an analysis of whole tissue blocks, the relative expression levels of tumor-elevated N-glycoforms are elevated in premalignant lesions compared to non-tumor

tissue. This finding suggests that aberrant N-glycosylation occurs during the process of tumor initiation, potentially implicating N-glycosylation in the tumorigenic process. This finding is clinically relevant, as early stage disease markers for pancreatic cancer should identify cancer prior to malignancy.

8. In an analysis of whole tissue blocks, the relative expression levels of tumor-elevated N-glycoforms are also elevated in tumor tissue compared to benign, diseased tissue. This finding is important as CA19-9 does not adequately distinguish tumor and benign diseases, leading to a poor specificity.

6.2.2 Limitations and Future Research

While this aim identified tumor associated N-glycans, there are several limitations that should be addressed to improve the study. The primary limitations for this aim can be categorized into [1] MALDI-IMS limitations, [2] data processing, and [3] disease specific limitations.

For more information on the limitations in MALDI-IMS for glycomic analysis, a thorough review is present in section 6.1.2. The main limitations that pertain to this aim involve comparative studies across multiple imaging experiments, the labile nature of sialic acid, and complications determining the anomeric linkages of monosaccharides. In pancreatic cancer, the elevation of the CA19-9 motif is well characterized, and consists of a terminal α2,3-linked sialic acid and an outer arm fucose residue. In our analysis, few N-glycans are both sialylated and fucosylated, suggesting CA19-9 is absent on N-glycans. However, of the glycans that were elevated in pancreatic cancer, the top five in terms of log2FC possess a fucose residue. Determining the location and linkage of this fucose is important, as sialic acids are known to dissociate during MALDI ionization. Approaches to stabilize sialic acid residues during MALDI-IMS were outline in Section 6.1.2. However, the limitation of the unknown anomeric fucose residue can be resolved by other experimental techniques. Comunale et al. used HPLC analysis and fluorescence detection of labeled glycans in conjunction with sequential exoglycosidase

digestion to differentiate core from outer arm fucose residues on a protein of interest (453). Lectins with specificity for core fucose can also be used to enrich glycoproteins prior to digestion with PNGaseF and glycomic analysis (485). The absence of glycoforms of interest would suggest the presence of branched fucose residues. Methods that directly correlate with MALDI-IMS could also be applied to differentiate core vs. outer arm fucose residues. Digestion with fucosidases that are specific to core fucose, either before or after PNGaseF but prior to MALDI-IMS, should result in a mass shift from the loss of core fucose, if present. If the distribution of the glycoforms of interest remains the same after digestion with the fucosidase, the tumor glycans possess outer arm fucose. Alternatively, serial tissue sections could be stained with lectins with specificity for defined fucose linkage. However, this method would not directly confirm whether the presence of core or outer arm fucose is on N-glycans.

As noted in the text, data processing was observed to be a limitation for this aim. While peak picking identified 50 ions that were present across all six TMAs, 23 of which corresponded to N-glycans, there was a noticeable lack of some of the more prevalent N-glycans, such as Hex5HexNAc4. A manual inspection of an N-glycan library identified 50 N-glycans that were present in all six TMAs. Thus, during peak picking, a little less than half of the glycans that are present were picked. This deficiency further impacts biomarker panel generation, as the machine learning algorithms process fewer N-glycan signatures. Additionally, as outlined in Section 6.1.2, identified peaks were selected using mass spectral binning, which limits the high mass resolution achievable by the FT-MS (481). The process of mass spectral binning across six TMAs likely contributes to the artificially high PPM errors observed and may create slight differences across TMAs resulting in the exclusion of N-glycans from being identified by peak picking. Future directions dealing with MALDI-IMS of peak picking, mass accuracy, and analysis of multiple slides is outlined in Section 6.1.2. A novel direction related to this aim deals with the missing N-glycans from peak picking. The peak picking process is initiated in FlexImaging by performing

hierarchical clustering, which produces a bucket table of ions and the relative intensity of the ion at that location. Glycans that are not reported in the pick peaking data, but are manually observed in the data, are often missing from the bucket table for at least one TMA. To adjust for these omissions, the relative abundance of individual ions can be selected in FlexImaging and exported individually, enabling a complete N-glycan analysis by downstream processing.

Finally, this aim looks at glycan biomarkers for pancreatic cancer in general, as opposed to identifying unique differences between the subtypes of pancreatic cancer. While adapting the method to enable analysis of FFPE tissue blocks improves sample numbers for distinguishing tumor vs. non-tumor, PDAC makes up about 90% of all pancreatic cancer incidences, so sample constraints may prevent PDAC to be distinguished from other subtypes of pancreatic cancer at a statistically significant level.

Glycan biomarker panels and their ability to differentiate malignant from benign tissues, as well as identify premalignant lesions, should be investigated using a larger cohort of tissue specimen directed at those specific comparisons. The poor ability to distinguish malignant from benign conditions is a significant limitation of disease markers, not limited to CA19-9 for pancreatic cancer, at it lowers the predictive value positive of disease markers by falsely classifying individuals with benign diseases as having malignant diseases. Preliminary evidence from MALDI-IMS of pancreatic tissue blocks in conjunction with histopathology analysis, suggests that N-glycan patterns can differentiate tumor tissue from benign non-tumor tissue. This trend needs to be followed up in a larger data set to confirm this finding across multiple patients, preferably by conducting MALDI-IMS on TMAs directly constructed to compare malignant and benign pancreas tissue. In an analysis of one TMA that meet the criteria, Hex6dHex1HexNAc5 was more abundant in the tumor cores compared to chronic pancreatitis (Figure 49).

Pancreatic cancer is often too advanced at the time of diagnosis to allow for surgical resection. As such, biomarkers for pancreatic cancer should identify the cancer at early stages,

preferably in the premalignant (PanIN) stage. Preliminary evidence shows that glycan biomarker panels not only distinguish tumor from non-tumor tissue, but also premalignant lesions from non-tumor tissue. Efforts to investigate this trend across a larger tissue cohort should be investigated. However, this study may not be possible using human tissue specimens, as the disease detection is made in later stages of pancreatic cancer. To compensate for limitations in clinical specimen, mouse models of PDAC could be used for this analysis. Inducible *KRAS* mutations in mouse models result in the formation of early stage precancerous lesions, but do not result in PDAC. The combination of other genomic mutations with *KRAS* mutations are able to form PanIN lesions which progress to PDAC (486). Variation in the N-glycan profile of PanIN lesions that either progress or do not progress to PDAC would determine whether glycomic analysis could identify which lesions would ultimately progress to cancer and warrant surgical resection. These findings could be validated by histopathology-guided analysis of PanIN lesions in human tissue specimen.

While MALDI-IMS is capable of distinguishing tumor from non-tumor tissue specimen, it is unlikely that MALDI-IMS will be implemented the clinic. The cost, experience, and time required for imaging analysis by high-resolution mass spectrometers is prohibitive for this application. Therefore, identified glycans should be assessed by other techniques that are used in the clinic. Serial sections of TMAs should be stained with broad specificity lectins directed at the tumor biomarker panel. Differential lectin staining towards tri- and tetraantennary glycans (phaseolus vulgaris lectin, L-PHA), fucosylated N-glycans (lens culinaris lectin, LCH, aleuria aurantia lectin, AAL), and high-mannose glycans (canavalia ensiformis lectin, ConA) may reflect trends observed for the individual glycan ions. Alternatively, N-glycan release from tissue sections, derivatization and analysis by HPLC-fluorescence detection or mass spectrometry using MRM techniques could be applied.

While the direct correlation between the tissue and the disease makes tissues useful for early biomarker discovery experiments, identified disease markers should be tested in serum. This

is especially relevant in pancreatic cancer, where screening should occur in asymptomatic patients in order to identify the disease when resection is still possible. For this purpose, identified glycan biomarker panels should be tested in serum samples, comparing the serum of non-diseased patients to that obtained from patients with pancreatic cancer or pancreatitis. This approach could be adapted to mass spectrometry or lectin arrays. Interestingly, prior studies along these lines have concluded that serum from PDAC patients exhibited trends of elevated fucosylation and N-glycan branching, suggesting that the biomarker panel could be applied in the analysis of serum or biological fluids (385). However, the serum proteome is complicated by a variety of proteins that are present in high abundance and may conceal aberrations related to the disease N-glycome. The glycomic analysis of individual proteins that are specific to the disease, such as AFP for liver cancer, may be more useful than an overall glycomic profiling of the serum.

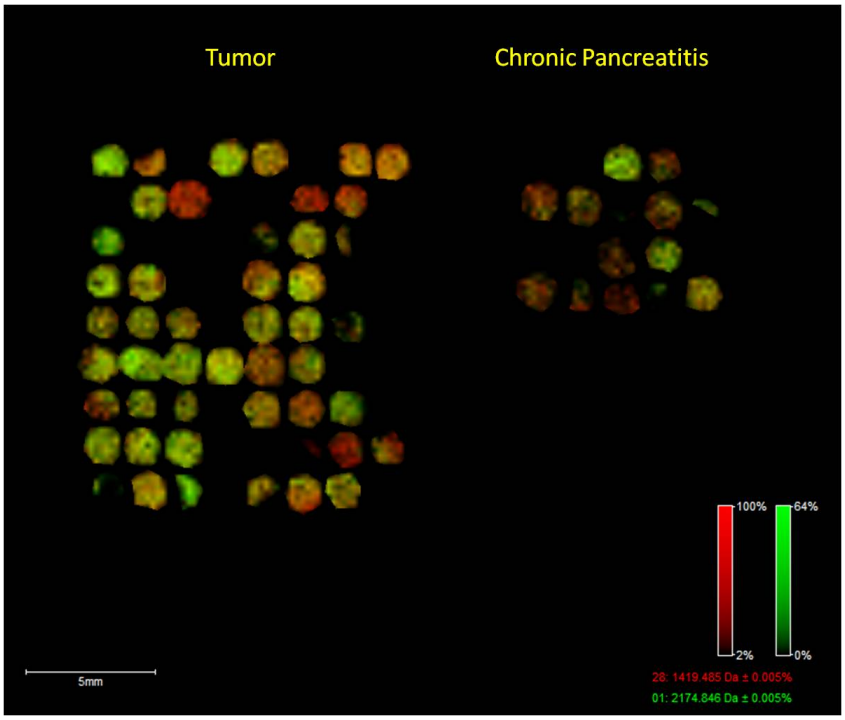


Figure 49. MALDI-IMS of a TMA Comparing Tumor and Pancreatitis. An overlay of Hex6dHex1HexNAc5 (green) and Hex6HexNAc2 (red), demonstrated that tumor cores can be differentiated from pancreatitis cores, by the same glycoforms that can distinguish tumor from non-diseased pancreas.

6.3 Glycoproteomic Analysis of Pancreatic Cancer

6.3.1 Conclusions

In an effort to identify proteins that exhibit altered glycosylation in pancreatic cancer, we performed a selective glycoprotein enrichment technique, as well as optimized a workflow for future intact glycopeptide analysis. We hypothesized that novel glycoproteomic studies would identify glycoproteins that differ in terms of protein abundance or protein glycosylation. A sialic acid analog was incorporated into four cell lines with documented CA19-9 expression levels to selectively enrich sialylated proteins and to assess the utility of the method in identifying CA19-9 carrier proteins. Furthermore, the role of N-glycoforms as CA19-9 carriers was examined by glycomic analysis and PNGaseF release of bound N-glycopeptides. A workflow for intact glycopeptide analysis, including enrichment strategies, mass spectrometry parameters, and data searching, was optimized. This method can be used to profile glycopeptides on selected proteins. By performing research relevant to this aim, we demonstrate:

1. N-linked glycans make little contribution to CA19-9 antigen expression. Glycomic analysis of high CA19-9 expressing cells reveal trends of elevated fucosylation, but no elevations in α2,3-linked sialic acids. The observation of elevated fucosylation suggests a role for the fucosyltransferases responsible for CA19-9 residues in N-glycosylation.
2. Click chemistry results in a highly efficient enrichment of sialylated glycoproteins. 80.2% of proteins identified are classified as glycoproteins in Uniprot. In total, 535 sialylated proteins were identified using click chemistry.
3. Overall elevations in total sialyltransferase levels do not seem to be present in high CA19-9 expressing cell lines. There is very little difference in the total number of proteins identified in the high CA19-9 expressing cell lines compared to the low CA19-9

cell lines (102 vs. 97). However, this does not exclude the possibility of elevations in specific sialyltransferases.

4. PNGaseF digestion of bound glycopeptides results in a small subset of identified glycoproteins, validating the glycomic profile results that suggest CA19-9 is primarily on O-glycans. Glycoproteins identified following PNGaseF release should be further assessed to determine if they are in fact carriers of the CA19-9 antigen

5. KEGG and GO pathway enrichment analyses reveal proteins related to O-glycan synthesis and processing to be among the most enriched pathways in the high CA19-9 cell lines

6. Click chemistry directed at sialylated glycoproteins identified numerous known CA19-9 carriers including members of the mucin and CEACAM family of proteins, as well as galectin-3-binding protein. Thus, the use of click chemistry on samples with defined CA19-9 status is an appropriate strategy to identify unknown CA19-9 carriers. However, identified proteins should be validated to ensure that the protein carries the CA19-9 antigen as opposed to just sialic acids.

7. For isolated glycoproteins, SAX and HILIC can effectively enrich glycopeptides over non-glycosylated peptides from the protein standards. In an analysis of fetuin and transferrin, all sites of N-glycosylation are observed in MALDI and LC-MS/MS analysis.

8. Intact glycopeptide analysis by HCD-pd-ETD of enriched glycopeptides demonstrates the concept of microheterogeneity and macroheterogeneity, both within and across glycoproteins. In the example provided, the number and composition of glycoforms differs with respect to individual peptides on the same glycoprotein.

9. A combined workflow integrating glycomic, proteomic, and glycopeptide analysis can generate a robust characterization of a glycoprotein or mixture of glycoproteins.

10. Selective enrichment of sialylated glycopeptides using hydrazide enrichment at low oxidizing conditions is consistent with glycopeptide enrichment using SAX or HILIC, with the exception that N-glycoforms no longer contain the sialic acid residue on the intact glycopeptide.

6.3.2 Limitations and Future Research

The initial goal of ManNAz-based click chemistry enrichment and downstream LC-MS/MS analysis was to identify putative carriers of the CA19-9 antigen and aberrantly expressed glycoproteins in cell lines. While this approach was validated by the consistency of this approach with published literature, there remain areas that can be further improved. While BXPC3 and Capan2 cell lines express high levels of CA19-9 and Panc1 and Miapaca2 express negligible levels of CA19-9, there are other intrinsic differences in the cell lines. Both BXPC3 and Capan2 cell lines are of epithelial morphology, while Panc1 and Miapaca2 exhibit a mesenchymal morphology. As multiple differences between the cell lines exist, some of the elevated glycoproteins may be a result of phenotypic differences not related to CA19-9 expression. These differences, as opposed to alterations in CA19-9 expression, likely contribute to E-cadherin elevations in the high CA19-9 cell lines and N-cadherin elevations in the low CA19-9 cell lines. Furthermore, neuropilin-1 and NCAM protein expression and sialylation have been linked to pancreatic cancer cell growth and invasion (487,488). Both proteins were expressed in higher abundance in the low CA19-9 cell lines. To adjust for variation in the characteristics of cell lines, most importantly differences in cellular morphology, additional controls can be implemented. Knockdown of EGFR in Panc1 cell lines resulted in a epithelial phenotype for the Panc1 cells and a concomitant decrease in N-cadherin and an increase in E-cadherin (489). As many proteins identified as differently upregulated in the low CA19-9 expressing cell lines are linked to cell migration and angiogenesis, using low CA19-9 expressing cells with an epithelial morphology may reduce identifications not attributed to CA19-9.

Sialylated glycoproteins were catalogued using click chemistry, but it is not known how sialylation of these proteins varies with respect to pancreatic cancer. No immortalized epithelial cell lines from normal human pancreatic duct cell (HPDE) lines were used in the experiments above. HPDE cell lines exhibit similarities in gene expression and morphology to normal pancreas cells (490). HPDE cells should be incorporated into the click chemistry study. While there is significant variation between high and low CA19-9 expressing cell lines, 337 proteins exhibited a $\leq$ 5-fold change in normalized spectral counts. Using an immortalized cell line that closely resembles normal pancreas ductal cells could identify differences in sialylated glycoproteins that are consistently elevated in the tumor cells. Sialylated glycoproteins identified uniquely in the tumor cells should be further investigated at disease markers for pancreatic cancer. This experiment, as opposed to the one implemented, would identify biomarker for pancreatic cancer in general, regardless of CA19-9 expression. Sialylated glycoproteins elevated in both high and low CA19-9 expressing cell lines compared to the HPDE cells could represent biomarker targets with higher sensitivity and specificity than CA19-9 alone. While no immortalized non-cancerous cell line was used, 127 sialylated glycoproteins were identified to be present in all four of the cell lines. Both *KRAS* and *TP53* were found to be significantly enriched upstream regulators by Ingenuity Pathway Analysis (IPA), based off of the large number of downstream proteins in the dataset consisting of the 127 glycoproteins found in all four of the cell lines (Figure 50). Furthermore, the biological process of tumor growth and migration of tumor cell lines were significantly enriched, and many of the proteins involved in these processes are downstream of *KRAS and TP53* (Figure 51. Interestingly, proteins downstream of *KRAS* (red boxes) appear to play a significant role in tumor growth, while those downstream of *TP53* (green boxes) play a significant role in migration. This is consistant with the genetic mutations that occur during tumorigenisis, where *KRAS* plays and itiating role in tumor progression but secondary

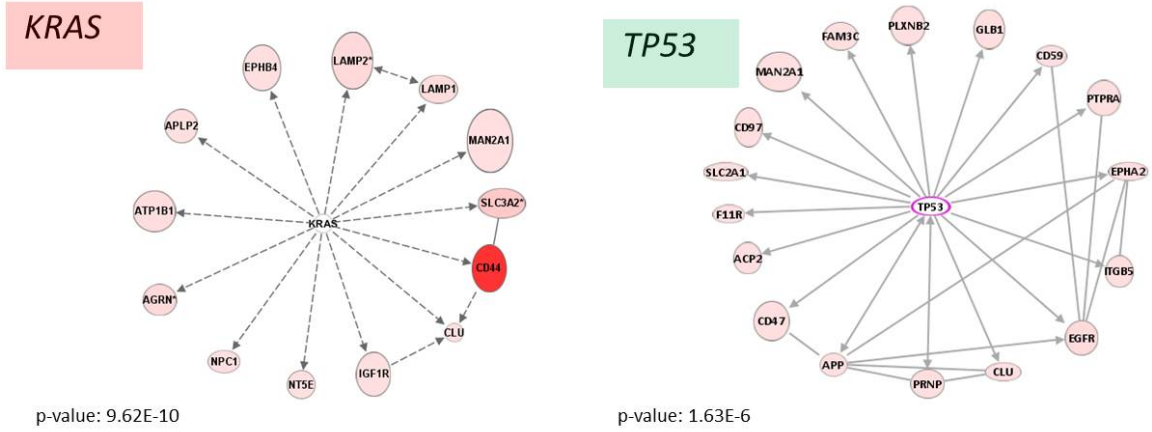


Figure 50. Ingenuity Pathways Analysis Reveals Important Upstream Regulators. The 127 proteins observed in across all four cell lines were searched with Ingenuity Pathway Analysis and identified several enriched upstream regulators, including *KRAS* and *TP53*.

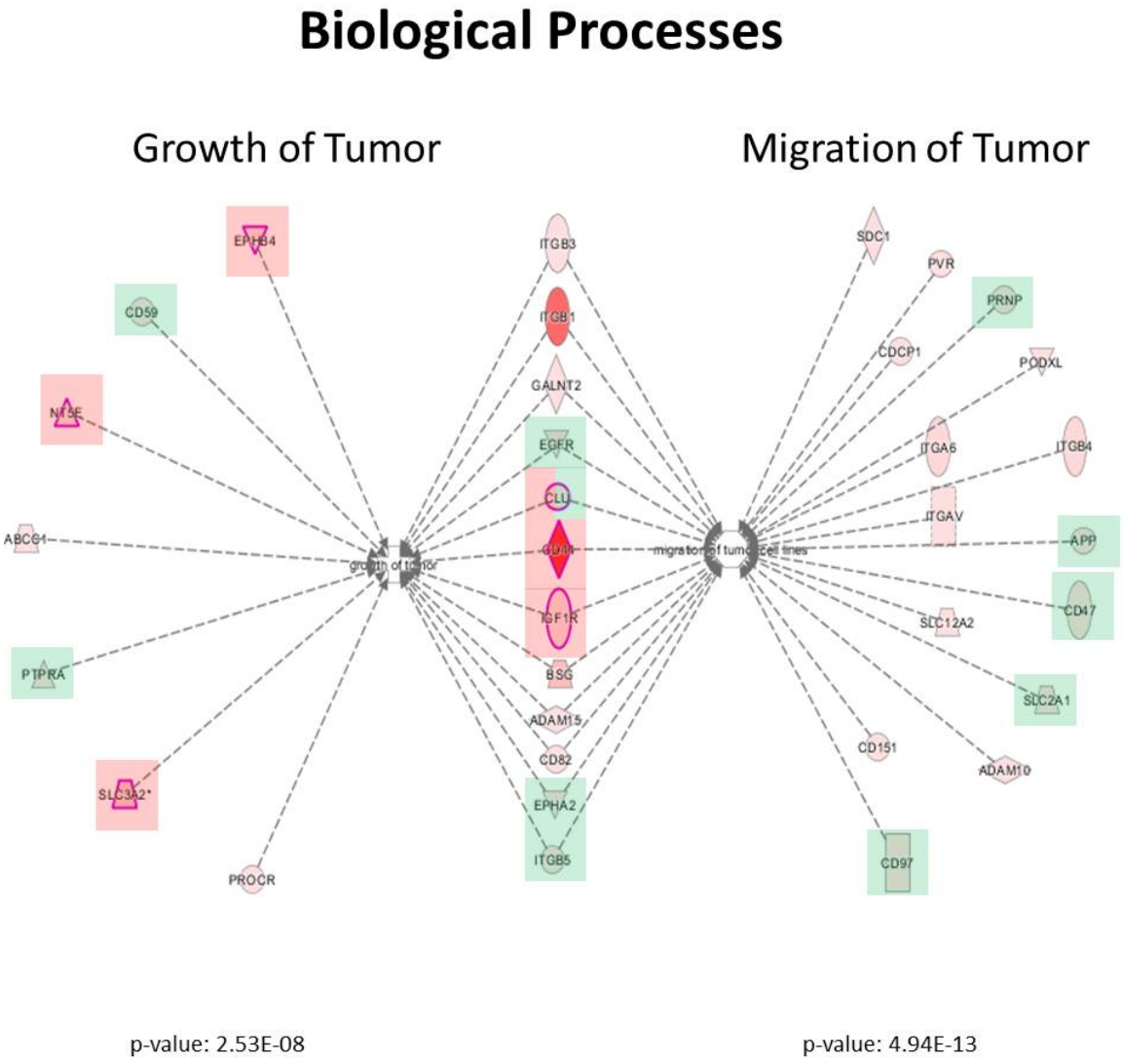


Figure 51. Ingenuity Pathway Analysis Reveals Several Biological Processes Involved in Pancreatic Cancer. Processes related to tumor growth and migration are heavily enriched in the glycoproteins identified in all four cell lines. Proteins that play a role in the processes are downstream of both *KRAS* (red box) and *TP53* (green box) are involved in these processes.

processes are required to initiate metastatic and aggressive phenotypes. Furthermore, this analysis identified several biomarker targets for pancreatic cancer that can be confirmed by additional screening approaches in biological specimen. These targets include clusterin and CD44, which have been previously documented in pancreatic cancer as playing a role in drug resistance and aggressiveness.

As a technique, click chemistry enrichment of glycoproteins has several limitations. Glycoproteins are bound to the alkyne-agarose resin through the incorporated sugar analog. Glycoprotein identification by mass spectrometry is based on tryptic peptides from glycoproteins that are not glycosylated, while glycopeptides remain bound to the resin. Thus, the actual site of glycosylation is not present in the LC-MS/MS data. PNGaseF release of bound N-glycopeptides enables identification of previously N-glycosylated peptides, but glycoform specific information is not obtainable. The use of the ManNAz label in the study limited analysis to sialylated glycopeptides. Based off of glycomic information, it is likely that there are glycoprotein differences that are not linked to sialic acid. Finally, while elevations in sialylated glycoproteins were reported, it is unclear whether the differences were at the level of protein expression or glycoform structure/occupancy. Glycomic analysis revealed significant trends in aberrant in N-glycan fucosylation that should be analyzed. Incorporation of fucose analogs and subsequent click chemistry can interrogate differences in fucosylation as a result of CA19-9 expression (289,302). Quantitative mass spectrometry at the proteomic level can identify whether the changes observed are the glycoform level or glycoprotein level. Stable isotope labeling by amino acids in cell culture (SILAC) can be implemented in conjunction with click chemistry. Alterations at the protein level should be reflected in the SILAC experiments, while alterations at the glycoform level will not appear in the SILAC data. Alternatively, Western blotting of selected proteins from the click chemistry data can be used to determine changes in protein expression.

Click chemistry cannot be used for intact glycopeptide analysis or analysis of global glycosylation. To obtain this information, selected glycoproteins should be subjected to the Orbitrap workflow and data processing outlined in the dissertation. As referenced, additional optimization is required to enable analysis of complex protein mixtures by intact glycopeptide analysis. However, it has been suggested that it is more beneficial to target individual proteins, as global examination cannot adequately profile the complexity of the proteome to include changes in glycosylation. Intact glycopeptide analysis using SAX failed to enrich glycopeptides from asialofetuin, suggesting that only sialylated glycopeptides can be enriched using this technique. Both SAX and HILIC failed to enrich glycopeptides from complex mixtures adequately enough to enable glycopeptide profiling. In order to overcome issues of sample complexity, Parker et al. instituted a combinatorial enrichment strategy involving HILIC enrichment and RPLC fractionation, prior to online LC-MS/MS analysis (304). Thus, it is unlikely that a single enrichment approach can be used for complex samples. In addition to enrichment of glycoproteins using physiochemical enrichment strategies, lectin enrichment of complex fucosylated N-glycans could identify numerous differences in intact N-glycoprotein analysis.

REFERENCES

1. Jemal A, Simard EP, Dorell C, Noone A-M, Markowitz LE, Kohler B, et al. Annual Report to the Nation on the Status of Cancer, 1975-2009, Featuring the Burden and Trends in Human Papillomavirus (HPV)-Associated Cancers and HPV Vaccination Coverage Levels. JNCI J Natl Cancer Inst. 2013 Feb 6;105(3):175–201.

2. Ma J, Jemal A. The rise and fall of cancer mortality in the USA: why does pancreatic cancer not follow the trend? Future Oncol Lond Engl. 2013 Jul;9(7):917–9.

3. Wolfgang CL, Corl F, Johnson PT, Edil BH, Horton KM, Schulick RD, et al. Pancreatic surgery for the radiologist, 2011: an illustrated review of classic and newer surgical techniques for pancreatic tumor resection. AJR Am J Roentgenol. 2011 Dec;197(6):1343–50.

4. Signs and Symptoms of Pancreatic Cancer [Internet]. American Cancer Society. 2014 [cited 2015 Apr 14]. Available from: http://www.cancer.org/cancer/pancreaticcancer/detailedguide/pancreatic-cancer-signs-and-symptoms

5. SEER Stat Fact Sheets: Pancreatic Cancer [Internet]. National Cancer Institute: Surveillance, Epidemiology, and End Results Program. [cited 2015 Apr 14]. Available from: http://seer.cancer.gov/statfacts/html/pancreas.html

6. Pancreatic Cancer Treatment [Internet]. National Cancer Institute. 2015 [cited 2015 Apr 14]. Available from: http://www.cancer.gov/cancertopics/pdq/treatment/pancreatic/Patient/page2

7. Pancreatic Cancer Survival by Stage [Internet]. American Cancer Society. 2015 [cited 2015 Apr 14]. Available from: http://www.cancer.org/cancer/pancreaticcancer/detailedguide/pancreatic-cancer-survival-rates

8. Al Haddad AHI, Adrian TE. Challenges and future directions in therapeutics for pancreatic ductal adenocarcinoma. Expert Opin Investig Drugs. 2014 Nov;23(11):1499–515.

9. Neesse A, Krug S, Gress TM, Tuveson DA, Michl P. Emerging concepts in pancreatic cancer medicine: targeting the tumor stroma. OncoTargets Ther. 2013;7:33–43.

10. Conroy T, Desseigne F, Ychou M, Bouché O, Guimbaud R, Bécouarn Y, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med. 2011 May 12;364(19):1817–25.

11. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. N Engl J Med. 2013 Oct 31;369(18):1691–703.

12. World Health Organization, International Agency for Research on Cancer. WHO classification of tumours of the digestive system. 4th ed. Bosman FT, editor. Lyon: International Agency for Research on Cancer; 2010. 417 p.

13. Fesinmeyer MD. Differences in Survival by Histologic Type of Pancreatic Cancer. Cancer Epidemiol Biomarkers Prev. 2005 Jul 1;14(7):1766–73.

14. Hezel AF. Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev. 2006 May 15;20(10):1218–49.

15. Koorstra J-BM, Feldmann G, Habbe N, Maitra A. Morphogenesis of pancreatic cancer: role of pancreatic intraepithelial neoplasia (PanINs). Langenbecks Arch Surg. 2008 Jul;393(4):561–70.

16. Watson L. Exocrine Insufficiency and Pancreatic Enzyme Replacement Therapy in Pancreatic Cancer. Clin Oncol. 2010 Jun;22(5):391.

17. Kato H, Nakao A, Kishimoto W, Nonami T, Harada A, Hayakawa T, et al. 13C-labeled trioctanoin breath test for exocrine pancreatic function test in patients after pancreatoduodenectomy. Am J Gastroenterol. 1993 Jan;88(1):64–9.

18. D'Haese JG, Ceyhan GO, Demir IE, Layer P, Uhl W, Löhr M, et al. Pancreatic enzyme replacement therapy in patients with exocrine pancreatic insufficiency due to chronic pancreatitis: a 1-year disease management study on symptom control and quality of life. Pancreas. 2014 Aug;43(6):834–41.

19. Ro C, Chai W, Yu VE, Yu R. Pancreatic neuroendocrine tumors: biology, diagnosis,and treatment. Chin J Cancer. 2013 Jun;32(6):312–24.

20. Grozinsky-Glasberg S, Mazeh H, Gross DJ. Clinical features of pancreatic neuroendocrine tumors. J Hepato-Biliary-Pancreat Sci. 2015 Feb.

21. Lowenfels AB, Maisonneuve P. Epidemiology and risk factors for pancreatic cancer. Best Pract Res Clin Gastroenterol. 2006 Apr;20(2):197–209.

22. Fuchs CS, Colditz GA, Stampfer MJ, Giovannucci EL, Hunter DJ, Rimm EB, et al. A prospective study of cigarette smoking and the risk of pancreatic cancer. Arch Intern Med. 1996 Oct 28;156(19):2255–60.

23. Everhart J, Wright D. Diabetes mellitus as a risk factor for pancreatic cancer. A meta-analysis. JAMA J Am Med Assoc. 1995 May 24;273(20):1605–9.

24. Michaud DS, Giovannucci E, Willett WC, Colditz GA, Stampfer MJ, Fuchs CS. Physical activity, obesity, height, and the risk of pancreatic cancer. JAMA J Am Med Assoc. 2001 Aug 22;286(8):921–9.

25. Berrington de Gonzalez A, Sweetland S, Spencer E. A meta-analysis of obesity and the risk of pancreatic cancer. Br J Cancer. 2003 Aug 4;89(3):519–23.

26. Mauro LA, Herman JM, Jaffee EM, Laheru DA. Carcinoma of the Pancreas. Abeloff's Clinical Oncology. 5th Edition. Churchill Livingstone; p. 1397–415.

27. Genkinger JM, Spiegelman D, Anderson KE, Bergkvist L, Bernstein L, van den Brandt PA, et al. Alcohol intake and pancreatic cancer risk: a pooled analysis of fourteen cohort studies. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2009 Mar;18(3):765–76.

28. Duell EJ. Epidemiology and potential mechanisms of tobacco smoking and heavy alcohol consumption in pancreatic cancer. Mol Carcinog. 2012 Jan;51(1):40–52.

29. Huxley R, Ansary-Moghaddam A, Berrington de González A, Barzi F, Woodward M. Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. Br J Cancer. 2005 Jun 6;92(11):2076–83.

30. Schenk M, Schwartz AG, O'Neal E, Kinnard M, Greenson JK, Fryzek JP, et al. Familial risk of pancreatic cancer. J Natl Cancer Inst. 2001 Apr 18;93(8):640–4.

31. Klein AP, Brune KA, Petersen GM, Goggins M, Tersmette AC, Offerhaus GJA, et al. Prospective risk of pancreatic cancer in familial pancreatic cancer kindreds. Cancer Res. 2004 Apr 1;64(7):2634–8.

32. Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. The Lancet. 2011 Aug;378(9791):607–20.

33. Thompson D. Cancer Incidence in BRCA1 Mutation Carriers. CancerSpectrum Knowl Environ. 2002 Sep 18;94(18):1358–65.

34. Jones S, Hruban RH, Kamiyama M, Borges M, Zhang X, Parsons DW, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. Science. 2009 Apr 10;324(5924):217.

35. Lal G, Liu G, Schmocker B, Kaurah P, Ozcelik H, Narod SA, et al. Inherited predisposition to pancreatic adenocarcinoma: role of family history and germ-line p16, BRCA1, and BRCA2 mutations. Cancer Res. 2000 Jan 15;60(2):409–16.

36. Whitcomb DC, Gorry MC, Preston RA, Furey W, Sossenheimer MJ, Ulrich CD, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. Nat Genet. 1996 Oct;14(2):141–5.

37. Jancík S, Drábek J, Radzioch D, Hajdúch M. Clinical Relevance of KRAS in Human Cancers. J Biomed Biotechnol. 2010; 2010:1–13.

38. Bardeesy N, DePinho RA. Pancreatic cancer biology and genetics. Nat Rev Cancer. 2002 Dec;2(12):897–909.

39. Zorde Khvalevsky E, Gabai R, Rachmut IH, Horwitz E, Brunschwig Z, Orbach A, et al. Mutant KRAS is a druggable target for pancreatic cancer. Proc Natl Acad Sci. 2013 Dec 17;110(51):20723–8.

40. Bryant KL, Mancias JD, Kimmelman AC, Der CJ. KRAS: feeding pancreatic cancer proliferation. Trends Biochem Sci. 2014 Feb;39(2):91–100.

41. Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D. RAS oncogenes: weaving a tumorigenic web. Nat Rev Cancer. 2011 Oct 13;11(11):761–74.

42. Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. Oncogene. 2007 May 14;26(22):3279–90.

43. Collisson EA, Trejo CL, Silva JM, Gu S, Korkola JE, Heiser LM, et al. A central role for RAF→MEK→ERK signaling in the genesis of pancreatic ductal adenocarcinoma. Cancer Discov. 2012 Aug;2(8):685–93.

44. Asano T, Yao Y, Zhu J, Li D, Abbruzzese JL, Reddy SAG. The PI 3-kinase/Akt signaling pathway is activated due to aberrant Pten expression and targets transcription factors NF-kappaB and c-Myc in pancreatic cancer cells. Oncogene. 2004 Nov 11;23(53):8571–80.

45. Cheng JQ, Ruggeri B, Klein WM, Sonoda G, Altomare DA, Watson DK, et al. Amplification of AKT2 in human pancreatic cells and inhibition of AKT2 expression and tumorigenicity by antisense RNA. Proc Natl Acad Sci U S A. 1996 Apr 16;93(8):3636–41.

46. Eser S, Reiff N, Messer M, Seidler B, Gottschalk K, Dobler M, et al. Selective Requirement of PI3K/PDK1 Signaling for Kras Oncogene-Driven Pancreatic Cell Plasticity and Cancer. Cancer Cell. 2013 Mar;23(3):406–20.

47. Hirano T, Shino Y, Saito T, Komoda F, Okutomi Y, Takeda A, et al. Dominant negative MEKK1 inhibits survival of pancreatic cancer cells. Oncogene. 2002 Aug 29;21(38):5923–8.

48. Asano T, Yao Y, Zhu J, Li D, Abbruzzese JL, Reddy SA. The rapamycin analog CCI-779 is a potent inhibitor of pancreatic cancer cell proliferation. Biochem Biophys Res Commun. 2005 May 27;331(1):295–302.

49. Hustinx SR, Leoni LM, Yeo CJ, Brown PN, Goggins M, Kern SE, et al. Concordant loss of MTAP and p16/CDKN2A expression in pancreatic intraepithelial neoplasia: evidence of homozygous deletion in a noninvasive precursor lesion. Mod Pathol Off J U S Can Acad Pathol Inc. 2005 Jul 18(7):959–63.

50. Agarwal P, Lutful Kabir FM, DeInnocentes P, Curtis R. Tumor Suppressor Gene p16/INK4A/CDKN2A and Its Role in Cell Cycle Exit, Differentiation, and Determination of Cell Fate. In: Cheng Y, editor. Tumor Suppressor Genes [Internet]. InTech; 2012 [cited 2014 Sep 2]. Available from: http://www.intechopen.com/books/tumor-suppressor-genes/tumor-suppressor-gene-p16-ink4a-cdkn2a-and-its-role-in-cell-cycle-exit-differentiation-and-determina

51. Petitjean A, Achatz MIW, Borresen-Dale AL, Hainaut P, Olivier M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. Oncogene. 2007 Apr 2;26(15):2157–65.

52. Belda-Iniesta C, Ibáñez de Cáceres I, Barriuso J, Castro Carpeño J, González Barón M, Feliú J. Molecular biology of pancreatic cancer. Clin Transl Oncol. 2008 Sep;10(9):530–7.

53. Srivastava S, Wang S, Tong YA, Hao ZM, Chang EH. Dominant negative effect of a germ-line mutant p53: a step fostering tumorigenesis. Cancer Res. 1993 Oct 1;53(19):4452–5.

54. Redston MS, Caldas C, Seymour AB, Hruban RH, da Costa L, Yeo CJ, et al. p53 mutations in pancreatic carcinoma and evidence of common involvement of homocopolymer tracts in DNA microdeletions. Cancer Res. 1994 Jun 1;54(11):3025–33.

55. Goh AM, Coffill CR, Lane DP. The role of mutant p53 in human cancer. J Pathol. 2011 Jan;223(2):116–26.

56. Morton JP, Timpson P, Karim SA, Ridgway RA, Athineos D, Doyle B, et al. Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. Proc Natl Acad Sci U S A. 2010 Jan 5;107(1):246–51.

57. Levy L, Hill CS. Smad4 Dependency Defines Two Classes of Transforming Growth Factor (TGF-) Target Genes and Distinguishes TGF- -Induced Epithelial-Mesenchymal Transition from Its Antiproliferative and Migratory Responses. Mol Cell Biol. 2005 Sep 15;25(18):8108–25.

58. Lin HY, Wang XF, Ng-Eaton E, Weinberg RA, Lodish HF. Expression cloning of the TGF-beta type II receptor, a functional transmembrane serine/threonine kinase. Cell. 1992 Feb 21;68(4):775–85.

59. Ellenrieder V, Hendler SF, Boeck W, Seufferlein T, Menke A, Ruhland C, et al. Transforming growth factor beta1 treatment leads to an epithelial-mesenchymal transdifferentiation of pancreatic cancer cells requiring extracellular signal-regulated kinase 2 activation. Cancer Res. 2001 May 15;61(10):4222–8.

60. Ellenrieder V, Hendler SF, Ruhland C, Boeck W, Adler G, Gress TM. TGF-beta-induced invasiveness of pancreatic cancer cells is mediated by matrix metalloproteinase-2 and the urokinase plasminogen activator system. Int J Cancer J Int Cancer. 2001 Jul 15;93(2):204–11.

61. Su GH, Hruban RH, Bansal RK, Bova GS, Tang DJ, Shekher MC, et al. Germline and somatic mutations of the STK11/LKB1 Peutz-Jeghers gene in pancreatic and biliary cancers. Am J Pathol. 1999 Jun;154(6):1835–40.

62. Murphy KM, Brune KA, Griffin C, Sollenberger JE, Petersen GM, Bansal R, et al. Evaluation of candidate genes MAP2K4, MADH4, ACVR1B, and BRCA2 in familial pancreatic cancer: deleterious BRCA2 mutations in 17%. Cancer Res. 2002 Jul 1;62(13):3789–93.

63. Berven FS, Ahmad R, Clauser KR, Carr SA. Optimizing Performance of Glycopeptide Capture for Plasma Proteomics. J Proteome Res. 2010 Apr 5;9(4):1706–15.

64. Wei C, Amos CI, Zhang N, Wang X, Rashid A, Walker CL, et al. Suppression of Peutz-Jeghers polyposis by targeting mammalian target of rapamycin signaling. Clin Cancer Res Off J Am Assoc Cancer Res. 2008 Feb 15;14(4):1167–71.

65. Sclabas GM, Fujioka S, Schmidt C, Evans DB, Chiao PJ. NF-kappaB in pancreatic cancer. Int J Gastrointest Cancer. 2003;33(1):15–26.

66. Wang W, Abbruzzese JL, Evans DB, Larry L, Cleary KR, Chiao PJ. The nuclear factor-kappa B RelA transcription factor is constitutively activated in human pancreatic adenocarcinoma cells. Clin Cancer Res Off J Am Assoc Cancer Res. 1999 Jan;5(1):119–27.

67. Arlt A, Vorndamm J, Breitenbroich M, Fölsch UR, Kalthoff H, Schmidt WE, et al. Inhibition of NF-kappaB sensitizes human pancreatic carcinoma cells to apoptosis induced by etoposide (VP16) or doxorubicin. Oncogene. 2001 Feb 15;20(7):859–68.

68. Berman DM, Karhadkar SS, Maitra A, Montes De Oca R, Gerstenblith MR, Briggs K, et al. Widespread requirement for Hedgehog ligand stimulation in growth of digestive tract tumours. Nature. 2003 Oct 23;425(6960):846–51.

69. Thayer SP, di Magliano MP, Heiser PW, Nielsen CM, Roberts DJ, Lauwers GY, et al. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. Nature. 2003 Oct 23;425(6960):851–6.

70. Barton CM, Hall PA, Hughes CM, Gullick WJ, Lemoine NR. Transforming growth factor alpha and epidermal growth factor in human pancreatic cancer. J Pathol. 1991 Feb;163(2):111–6.

71. Friess H, Yamanaka Y, Kobrin MS, Do DA, Büchler MW, Korc M. Enhanced erbB-3 expression in human pancreatic cancer correlates with tumor progression. Clin Cancer Res Off J Am Assoc Cancer Res. 1995 Nov;1(11):1413–20.

72. Bergmann U, Funatomi H, Yokoyama M, Beger HG, Korc M. Insulin-like growth factor I overexpression in human pancreatic cancer: evidence for autocrine and paracrine roles. Cancer Res. 1995 May 15;55(10):2007–11.

73. Nair PN, De Armond DT, Adamo ML, Strodel WE, Freeman JW. Aberrant expression and activation of insulin-like growth factor-1 receptor (IGF-1R) are mediated by an induction of IGF-1R promoter activity and stabilization of IGF-1R mRNA and contributes to growth factor independence and increased survival of the pancreatic cancer cell line MIA PaCa-2. Oncogene. 2001 Dec 13;20(57):8203–14.

74. Itakura J, Ishiwata T, Friess H, Fujii H, Matsumoto Y, Büchler MW, et al. Enhanced expression of vascular endothelial growth factor in human pancreatic cancer correlates with local disease progression. Clin Cancer Res Off J Am Assoc Cancer Res. 1997 Aug;3(8):1309–16.

75. Li J, Kleeff J, Giese N, Büchler MW, Korc M, Friess H. Gefitinib (’Iressa’, ZD1839), a selective epidermal growth factor receptor tyrosine kinase inhibitor, inhibits pancreatic cancer cell growth, invasion, and colony formation. Int J Oncol. 2004 Jul;25(1):203–10.

76. Bruns CJ, Solorzano CC, Harbison MT, Ozawa S, Tsan R, Fan D, et al. Blockade of the epidermal growth factor receptor signaling by a novel tyrosine kinase inhibitor leads to apoptosis of endothelial cells and therapy of human pancreatic carcinoma. Cancer Res. 2000 Jun 1;60(11):2926–35.

77. Von Marschall Z, Cramer T, Höcker M, Burde R, Plath T, Schirner M, et al. De novo expression of vascular endothelial growth factor in human pancreatic cancer: evidence for an autocrine mitogenic loop. Gastroenterology. 2000 Nov;119(5):1358–72.

78. Tokunaga T, Abe Y, Tsuchida T, Hatanaka H, Oshika Y, Tomisawa M, et al. Ribozyme mediated cleavage of cell-associated isoform of vascular endothelial growth factor inhibits liver metastasis of a pancreatic cancer cell line. Int J Oncol. 2002 Nov;21(5):1027–32.

79. Hernández-Muñoz I, Skoudy A, Real FX, Navarro P. Pancreatic ductal adenocarcinoma: cellular origin, signaling pathways and stroma contribution. Pancreatol Off J Int Assoc Pancreatol IAP Al. 2008;8(4-5):462–9.

80. Collins MA, Pasca di Magliano M. Kras as a key oncogene and therapeutic target in pancreatic cancer. Front Physiol [Internet]. 2014 [cited 2014 Apr 21];4. Available from: http://www.frontiersin.org/Gastrointestinal_Sciences/10.3389/fphys.2013.00407/abstract

81. Löhr M, Klöppel G, Maisonneuve P, Lowenfels AB, Lüttges J. Frequency of K-ras Mutations in Pancreatic Intraductal Neoplasias Associated with Pancreatic Ductal Adenocarcinoma and Chronic Pancreatitis: A Meta-Analysis. Neoplasia. 2005 Jan;7(1):17–23.

82. Brembeck FH, Schreiber FS, Deramaudt TB, Craig L, Rhoades B, Swain G, et al. The mutant K-ras oncogene causes pancreatic periductal lymphocytic infiltration and gastric mucous neck cell hyperplasia in transgenic mice. Cancer Res. 2003 May 1;63(9):2005–9.

83. Guerra C, Schuhmacher AJ, Cañamero M, Grippo PJ, Verdaguer L, Pérez-Gallego L, et al. Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. Cancer Cell. 2007 Mar;11(3):291–302.

84. Wagner M, Greten FR, Weber CK, Koschnick S, Mattfeldt T, Deppert W, et al. A murine tumor progression model for pancreatic cancer recapitulating the genetic alterations of the human disease. Genes Dev. 2001 Feb 1;15(3):286–93.

85. Sandgren EP, Quaife CJ, Paulovich AG, Palmiter RD, Brinster RL. Pancreatic tumor pathogenesis reflects the causative genetic lesion. Proc Natl Acad Sci U S A. 1991 Jan 1;88(1):93–7.

86. Tuveson DA, Zhu L, Gopinathan A, Willis NA, Kachatrian L, Grochow R, et al. Mist1-KrasG12D knock-in mice develop mixed differentiation metastatic exocrine pancreatic carcinoma and hepatocellular carcinoma. Cancer Res. 2006 Jan 1;66(1):242–7.

87. Li C, Lee CJ, Simeone DM. Identification of human pancreatic cancer stem cells. Methods Mol Biol Clifton NJ. 2009;568:161–73.

88. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, et al. Identification of pancreatic cancer stem cells. Cancer Res. 2007 Feb 1;67(3):1030–7.

89. Hermann PC, Huber SL, Herrler T, Aicher A, Ellwart JW, Guba M, et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. Cell Stem Cell. 2007 Sep 13;1(3):313–23.

90. Seufferlein T, Bachet JB, Van Cutsem E, Rougier P, on behalf of the ESMO Guidelines Working Group. Pancreatic adenocarcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012 Oct 1;23(suppl 7):vii33–40.

91. Warshaw AL, Castillo CF. Pancreatic Carcinoma. N Engl J Med. 1992 Feb 13;326(7):455–65.

92. How is Pancreatic Cancer Diagnosed [Internet]. American Cancer Society. 2015 [cited 2015 Apr 14]. Available from: http://www.cancer.org/cancer/pancreaticcancer/detailedguide/pancreatic-cancer-survival-rates

93. Winter JM, Maitra A, Yeo CJ. Genetics and pathology of pancreatic cancer. HPB. 2006;8(5):324–36.

94. Rasheed ZA, Matsui W, Maitra A. Pathology of pancreatic stroma in PDAC. In: Grippo PJ, Munshi HG, editors. Pancreatic Cancer and Tumor Microenvironment [Internet]. Trivandrum (India): Transworld Research Network; 2012 [cited 2014 Apr 8]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK98933/

95. Hamada S, Masamune A, Shimosegawa T. Alteration of pancreatic cancer cell functions by tumor-stromal cell interaction. Front Physiol. 2013;4:318.

96. Hruban RH, Fukushima N. Pancreatic adenocarcinoma: update on the surgical pathology of carcinomas of ductal origin and PanINs. Mod Pathol. 2007 Feb;20(1s):S61–70.

97. Song J-W, Lee J-H. New Morphological Features for Grading Pancreatic Ductal Adenocarcinomas. BioMed Res Int. 2013;2013:1–25.

98. Dole M. Molecular Beams of Macroions. J Chem Phys. 1968;49(5):2240.

99. Fenn JB. Electrospray ionization mass spectrometry: How it all began. J Biomol Tech JBT. 2002 Sep;13(3):101–18.

100. Konermann L, Ahadi E, Rodriguez AD, Vahidi S. Unraveling the Mechanism of Electrospray Ionization. Anal Chem. 2013 Jan 2;85(1):2–9.

101. Banerjee S, Mazumdar S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int J Anal Chem. 2012;2012:1–40.

102. Wilm M. Principles of electrospray ionization. Mol Cell Proteomics MCP. 2011 Jul;10(7):M111.009407.

103. Woods AG, Sokolowska I, Wetie AGN, Wormwood K, Aslebagh R, Patel S, et al. Mass Spectrometry for Proteomics-Based Investigation. In: Woods AG, Darie CC, editors.

Advancements of Mass Spectrometry in Biomedical Research [Internet]. Cham: Springer International Publishing; 2014 [cited 2015 Mar 16]. p. 1–32. Available from: http://link.springer.com/10.1007/978-3-319-06068-2_1

104. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 1988 Oct 15;60(20):2299–301.

105. Lewis K, Suizdak G. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry in Peptide and Protein Analysis. Encyclopedia of Analytical Chemistry. 2000: 5880-5894.

106. Trimpin S, Inutan ED, Herath TN, McEwen CN. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Method for Selectively Producing Either Singly or Multiply Charged Molecular Ions. Anal Chem. 2010 Jan;82(1):11–5.

107. Domon B, Aebersold R. Mass spectrometry and protein analysis. Science. 2006 Apr 14;312(5771):212–7.

108. Yang Y, Zhang S, Howe K, Wilson DB, Moser F, Irwin D, et al. A comparison of nLC-ESI-MS/MS and nLC-MALDI-MS/MS for GeLC-based protein identification and iTRAQ-based shotgun quantitative proteomics. J Biomol Tech JBT. 2007 Sep;18(4):226–37.

109. Girolamo F, Lante I, Muraca M, Putignani L. The Role of Mass Spectrometry in the "Omics" Era. Curr Org Chem. 2013 Dec 31;17(23):2891–905.

110. Miller PE, Denton MB. The quadrupole mass filter: Basic operating concepts. J Chem Educ. 1986 Jul;63(7):617.

111. Guilhaus M. Principles and Instrumentation in Time-of-Fight Mass Spectrometry. J Mass Spec. 1995, Sept; 30: 1519-1532.

112. March R. An Introduction to Quadrupole Ion Trap Mass Spectrometry. J Mass Spectrom. J Mass Spec. 1998, Dec 4; 32(4): 351-369.

113. Schmid DG, Grosche P, Bandel H, Jung G. FTICR-mass spectrometry for high-resolution analysis in combinatorial chemistry. Biotechnol Bioeng. 2000 2001;71(2):149–61.

114. Payne AH, Glish GL. Tandem Mass Spectrometry in Quadrupole Ion Trap and Ion Cyclotron Resonance Mass Spectrometers. Methods in Enzymology [Internet]. Elsevier; 2005 [cited 2015 Mar 16]. p. 109–48. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0076687905020045

115. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci. 2004 Jun 29;101(26):9528–33.

116. Orbitrap Elite Hardware Manual. Thermo Fisher Scientific Inc.; 2011.

117. Shao C, Zhang Y, Sun W. Statistical characterization of HCD fragmentation patterns of tryptic peptides on an LTQ Orbitrap Velos mass spectrometer. J Proteomics. 2014 Sep;109:26–37.

118. Jungmann JH, Heeren RMA. Emerging technologies in mass spectrometry imaging. J Proteomics. 2012 Aug 30;75(16):5077–92.

119. Vestal ML. The future of biological mass spectrometry. J Am Soc Mass Spectrom. 2011 Jun;22(6):953–9.

120. Gross RW, Han X. Lipidomics at the interface of structure and function in systems biology. Chem Biol. 2011 Mar 25;18(3):284–91.

121. Han X, Gross RW. Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry: a bridge to lipidomics. J Lipid Res. 2003 Jun;44(6):1071–9.

122. Villas-Bôas SG, Mas S, Akesson M, Smedsgaard J, Nielsen J. Mass spectrometry in metabolome analysis. Mass Spectrom Rev. 2005 Oct;24(5):613–46.

123. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007 Feb;26(1):51–78.

124. Pernemalm M, Lehtiö J. Mass spectrometry-based plasma proteomics: state of the art and future outlook. Expert Rev Proteomics. 2014 Aug;11(4):431–48.

125. Bauer M, Ahrné E, Baron AP, Glatter T, Fava LL, Santamaria A, et al. Evaluation of Data-Dependent and -Independent Mass Spectrometric Workflows for Sensitive Quantification of Proteins and Phosphorylation Sites. J Proteome Res. 2014 Dec 5;13(12):5973–88.

126. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics MCP. 2012 Jun;11(6):O111.016717.

127. Doerr A. DIA mass spectrometry. Nat Methods. 2014 Dec 30;12(1):35–35.

128. Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR. Protein Analysis by Shotgun/Bottom-up Proteomics. Chem Rev. 2013 Apr 10;113(4):2343–94.

129. Yen T-Y, Dutta S, Litsakos-Cheung C, Corona A, Timpe L, Macher B. Overcoming Challenges and Opening New Opportunities in Glycoproteomics. Biomolecules. 2013 Mar 26;3(2):270–86.

130. Ladanyi A, Sipos F, Szoke D, Galamb O, Molnar B, Tulassay Z. Laser microdissection in translational and clinical research. Cytom Part J Int Soc Anal Cytol. 2006 Sep 1;69(9):947–60.

131. Mukherjee S, Rodriguez-Canales J, Hanson J, Emmert-Buck MR, Tangrea MA, Prieto DA, et al. Proteomic analysis of frozen tissue samples using laser capture microdissection. Methods Mol Biol Clifton NJ. 2013;1002:71–83.

132. Petyuk VA, Qian W-J, Chin MH, Wang H, Livesay EA, Monroe ME, et al. Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography-mass spectrometry. Genome Res. 2007 Mar;17(3):328–36.

133. BarcelÃ³-Coblijn G, FernÃ¡ndez JA. Mass spectrometry coupled to imaging techniques: the better the view the greater the challenge. Front Physiol [Internet]. 2015 Jan 22 [cited 2015 Feb 26];6. Available from: http://journal.frontiersin.org/article/10.3389/fphys.2015.00003/abstract

134. Winograd N, Bloom A. Sample preparation for 3D SIMS chemical imaging of cells. Methods Mol Biol Clifton NJ. 2015;1203:9–19.

135. Groseclose MR, Andersson M, Hardesty WM, Caprioli RM. Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry. J Mass Spectrom JMS. 2007 Feb;42(2):254–62.

136. Thomas A, Charbonneau JL, Fournaise E, Chaurand P. Sublimation of New Matrix Candidates for High Spatial Resolution Imaging Mass Spectrometry of Lipids: Enhanced Information in Both Positive and Negative Polarities after 1,5-Diaminonapthalene Deposition. Anal Chem. 2012 Feb 21;84(4):2048–54.

137. Puolitaival SM, Burnum KE, Cornett DS, Caprioli RM. Solvent-free matrix dry-coating for MALDI imaging of phospholipids. J Am Soc Mass Spectrom. 2008 Jun;19(6):882–6.

138. Chaurand P, Norris JL, Cornett DS, Mobley JA, Caprioli RM. New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. J Proteome Res. 2006 Nov;5(11):2889–900.

139. Stoeckli M, Chaurand P, Hallahan DE, Caprioli RM. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. Nat Med. 2001 Apr;7(4):493–6.

140. Chaurand P, Caprioli RM. Direct profiling and imaging of peptides and proteins from mammalian cells and tissue sections by mass spectrometry. Electrophoresis. 2002 Sep;23(18):3125–35.

141. Reyzer ML, Hsieh Y, Ng K, Korfmacher WA, Caprioli RM. Direct analysis of drug candidates in tissue by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom JMS. 2003 Oct;38(10):1081–92.

142. Hsieh Y, Casale R, Fukuda E, Chen J, Knemeyer I, Wingate J, et al. Matrix-assisted laser desorption/ionization imaging mass spectrometry for direct measurement of clozapine in rat brain tissue. Rapid Commun Mass Spectrom RCM. 2006;20(6):965–72.

143. Cornett DS, Frappier SL, Caprioli RM. MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. Anal Chem. 2008 Jul 15;80(14):5648–53.

144. Casadonte R, Caprioli RM. Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. Nat Protoc. 2011 Nov;6(11):1695–709.

145. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, et al. Mass spectrometry-based proteomics: existing capabilities and future directions. Chem Soc Rev. 2012 May 21;41(10):3912–28.

146. Eriksson C, Masaki N, Yao I, Hayasaka T, Setou M. MALDI Imaging Mass Spectrometry—A Mini Review of Methods and Recent Developments. Mass Spectrom. 2013;2(Special_Issue):S0022–S0022.

147. Aerni H-R, Cornett DS, Caprioli RM. Automated acoustic matrix deposition for MALDI sample preparation. Anal Chem. 2006 Feb 1;78(3):827–34.

148. SolariX User Manual. Bruker Daltonik GmbH; 2011.

149. Pachl F, Ruprecht B, Lemeer S, Kuster B. Characterization of a high field Orbitrap mass spectrometer for proteome analysis. Proteomics. 2013 Sep;13(17):2552–62.

150. Varki AA, Cummings RDRD, Esko JDJD, Freeze HHHH, Stanley PP, Bertozzi CRCR, et al., editors. Essentials of Glycobiology [Internet]. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009 [cited 2012 Aug 7]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20301279

151. Nguema-Ona E, Vicré-Gibouin M, Gotté M, Plancot B, Lerouge P, Bardor M, et al. Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. Front Plant Sci. 2014;5:499.

152. Kudelka MR, Ju T, Heimburg-Molinaro J, Cummings RD. Simple sugars to complex disease-mucin-type o-glycans in cancer. Adv Cancer Res. 2015;126:53–135.

153. Lowe JB, Marth JD. A genetic approach to Mammalian glycan function. Annu Rev Biochem. 2003;72:643–91.

154. Priatel JJ, Chui D, Hiraoka N, Simmons CJ, Richardson KB, Page DM, et al. The ST3Gal-I sialyltransferase controls CD8+ T lymphocyte homeostasis by modulating O-glycan biosynthesis. Immunity. 2000 Mar;12(3):273–83.

155. Prien JM, Huysentruyt LC, Ashline DJ, Lapadula AJ, Seyfried TN, Reinhold VN. Differentiating N-linked glycan structural isomers in metastatic and nonmetastatic tumor cells using sequential mass spectrometry. Glycobiology. 2008 May;18(5):353–66.

156. Ruhaak LR, Miyamoto S, Lebrilla CB. Developments in the identification of glycan biomarkers for the detection of cancer. Mol Cell Proteomics MCP. 2013 Apr;12(4):846–55.

157. National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. Transforming Glycoscience: A Roadmap for the Future [Internet]. Washington (DC): National Academies Press (US); 2012 [cited 2014 Jul 30]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK109958/

158. Pang P-C, Chiu PCN, Lee C-L, Chang L-Y, Panico M, Morris HR, et al. Human sperm binding is mediated by the sialyl-Lewis(x) oligosaccharide on the zona pellucida. Science. 2011 Sep 23;333(6050):1761–4.

159. Leppänen A, Mehta P, Ouyang YB, Ju T, Helin J, Moore KL, et al. A novel glycosulfopeptide binds to P-selectin and inhibits leukocyte adhesion to P-selectin. J Biol Chem. 1999 Aug 27;274(35):24838–48.

160. Badr HA, Alsadek DMM, Darwish AA, Elsayed AI, Bekmanov BO, Khussainova EM, et al. Lectin approaches for glycoproteomics in FDA-approved cancer biomarkers. Expert Rev Proteomics. 2014 Apr;11(2):227–36.

161. Peracaula R, Barrabés S, Sarrats A, Rudd PM, de Llorens R. Altered glycosylation in tumours focused to cancer diagnosis. Dis Markers. 2008;25(4-5):207–18.

162. Milland J, Sandrin MS. ABO blood group and related antigens, natural antibodies and transplantation. Tissue Antigens. 2006 Dec;68(6):459–66.

163. Green RS, Stone EL, Tenno M, Lehtonen E, Farquhar MG, Marth JD. Mammalian N-glycan branching protects against innate immune self-recognition and inflammation in autoimmune disease pathogenesis. Immunity. 2007 Aug;27(2):308–20.

164. Goulabchand R, Vincent T, Batteux F, Eliaou J-F, Guilpain P. Impact of autoantibody glycosylation in autoimmune diseases. Autoimmun Rev. 2014 Jul;13(7):742–50.

165. Skålén K, Gustafsson M, Rydberg EK, Hultén LM, Wiklund O, Innerarity TL, et al. Subendothelial retention of atherogenic lipoproteins in early atherosclerosis. Nature. 2002 Jun 13;417(6890):750–4.

166. Bourdillon M-C, Randon J, Barek L, Zibara K, Covacho C, Poston RN, et al. Reduced Atherosclerotic Lesion Size in P -Selectin Deficient Apolipoprotein E -Knockout Mice Fed a Chow but Not a Fat Diet. J Biomed Biotechnol. 2006:1–8.

167. Dong ZM, Chapman SM, Brown AA, Frenette PS, Hynes RO, Wagner DD. The combined role of P- and E-selectins in atherosclerosis. J Clin Invest. 1998 Jul 1;102(1):145–52.

168. Gamblin SJ, Skehel JJ. Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins. J Biol Chem. 2010 Sep 10;285(37):28403–9.

169. MacKenzie CR, Hirama T, Lee KK, Altman E, Young NM. Quantitative analysis of bacterial toxin affinity and specificity for glycolipid receptors by surface plasmon resonance. J Biol Chem. 1997 Feb 28;272(9):5533–8.

170. Kawakubo M, Ito Y, Okimura Y, Kobayashi M, Sakura K, Kasama S, et al. Natural antibiotic function of a human gastric mucin against Helicobacter pylori infection. Science. 2004 Aug 13;305(5686):1003–6.

171. Wang W, Nie J, Prochnow C, Truong C, Jia Z, Wang S, et al. A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. Retrovirology. 2013;10:14.

172. Dalziel M, Crispin M, Scanlan CN, Zitzmann N, Dwek RA. Emerging Principles for the Therapeutic Exploitation of Glycosylation. Science. 2014 Jan 3;343(6166):1235681–1235681.

173. Buskas T, Thompson P, Boons G-J. Immunotherapy for cancer: synthetic carbohydrate-based vaccines. Chem Commun Camb Engl. 2009 Sep 28;(36):5335–49.

174. Ryan SO, Turner MS, Gariépy J, Finn OJ. Tumor antigen epitopes interpreted by the immune system as self or abnormal-self differentially affect cancer vaccine responses. Cancer Res. 2010 Jul 15;70(14):5788–96.

175. Lakshminarayanan V, Thompson P, Wolfert MA, Buskas T, Bradley JM, Pathangey LB, et al. Immune recognition of tumor-associated mucin MUC1 is achieved by a fully synthetic aberrantly glycosylated MUC1 tripartite vaccine. Proc Natl Acad Sci U S A. 2012 Jan 3;109(1):261–6.

176. Yu Y, Mishra S, Song X, Lasanajak Y, Bradley KC, Tappert MM, et al. Functional Glycomic Analysis of Human Milk Glycans Reveals the Presence of Virus Receptors and Embryonic Stem Cell Biomarkers. J Biol Chem. 2012 Dec 28;287(53):44784–99.

177. Peterson R, Cheah WY, Grinyer J, Packer N. Glycoconjugates in human milk: protecting infants from disease. Glycobiology. 2013 Dec;23(12):1425–38.

178. Lew W, Chen X, Kim CU. Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor. Curr Med Chem. 2000 Jun;7(6):663–72.

179. Von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature. 1993 Jun 3;363(6428):418–23.

180. Mizuochi T, Spellman MW, Larkin M, Solomon J, Basa LJ, Feizi T. Carbohydrate structures of the human-immunodeficiency-virus (HIV) recombinant envelope glycoprotein gp120 produced in Chinese-hamster ovary cells. Biochem J. 1988 Sep 1;254(2):599–603.

181. Lühn K, Wild MK. Human deficiencies of fucosylation and sialylation affecting selectin ligands. Semin Immunopathol. 2012 May;34(3):383–99.

182. Morel CF, Clarke JTR. The use of agalsidase alfa enzyme replacement therapy in the treatment of Fabry disease. Expert Opin Biol Ther. 2009 May;9(5):631–9.

183. Kakkis ED, Muenzer J, Tiller GE, Waber L, Belmont J, Passage M, et al. Enzyme-replacement therapy in mucopolysaccharidosis I. N Engl J Med. 2001 Jan 18;344(3):182–8.

184. Wang L-X, Lomino JV. Emerging technologies for making glycan-defined glycoproteins. ACS Chem Biol. 2012 Jan 20;7(1):110–22.

185. Solá RJ, Griebenow K. Glycosylation of therapeutic proteins: an effective strategy to optimize efficacy. BioDrugs Clin Immunother Biopharm Gene Ther. 2010 Feb 1;24(1):9–21.

186. Nyathi Y, Wilkinson BM, Pool MR. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Biochim Biophys Acta. 2013 Nov;1833(11):2392–402.

187. Lederkremer GZ. Glycoprotein folding, quality control and ER-associated degradation. Curr Opin Struct Biol. 2009 Oct;19(5):515–23.

188. Burda P, Aebi M. The dolichol pathway of N-linked glycosylation. Biochim Biophys Acta. 1999 Jan 6;1426(2):239–57.

189. Schwarz F, Aebi M. Mechanisms and principles of N-linked protein glycosylation. Curr Opin Struct Biol. 2011 Oct;21(5):576–82.

190. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT-BG, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 2013 May 15;32(10):1478–88.

191. Aryal RP, Ju T, Cummings RD. Identification of a novel protein binding motif within the T-synthase for the molecular chaperone Cosmc. J Biol Chem. 2014 Apr 25;289(17):11630–41.

192. Boyce M, Carrico IS, Ganguli AS, Yu S-H, Hangauer MJ, Hubbard SC, et al. Metabolic cross-talk allows labeling of O-linked beta-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. Proc Natl Acad Sci U S A. 2011 Feb 22;108(8):3141–6.

193. Ozanne B, Sambrook J. Binding of radioactively labelled concanavalin A and wheat germ agglutinin to normal and virus-transformed cells. Nature New Biol. 1971 Aug 4;232(31):156–60.

194. Aub JC, Tieslau C, Lankester A. Reactions of normal and tumor cell surfaces to enzymes. I. Wheat-germ lipase and associated mucopolysaccharides. Proc Natl Acad Sci U S A. 1963 Oct;50:613–9.

195. Magnani JL, Steplewski Z, Koprowski H, Ginsburg V. Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19-9 in the sera of patients as a mucin. Cancer Res. 1983 Nov;43(11):5489–92.

196. Nuti M, Teramoto YA, Mariani-Costantini R, Hand PH, Colcher D, Schlom J. A monoclonal antibody (B72.3) defines patterns of distribution of a novel tumor-associated

antigen in human mammary carcinoma cell populations. Int J Cancer J Int Cancer. 1982 May 15;29(5):539–45.

197. Yin BW, Lloyd KO. Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16. J Biol Chem. 2001 Jul 20;276(29):27371–5.

198. Biancone L, Araki M, Araki K, Vassalli P, Stamenkovic I. Redirection of tumor metastasis by expression of E-selectin in vivo. J Exp Med. 1996 Feb 1;183(2):581–7.

199. Feizi T. Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens. Nature. 1985 Mar 7;314(6006):53–7.

200. Springer GF. Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. J Mol Med Berl Ger. 1997 Aug;75(8):594–602.

201. Desai PR. Immunoreactive T and Tn antigens in malignancy: role in carcinoma diagnosis, prognosis, and immunotherapy. Transfus Med Rev. 2000 Oct;14(4):312–25.

202. Pinho SS, Oliveira P, Cabral J, Carvalho S, Huntsman D, Gärtner F, et al. Loss and recovery of Mgat3 and GnT-III Mediated E-cadherin N-glycosylation is a mechanism involved in epithelial-mesenchymal-epithelial transitions. PloS One. 2012;7(3):e33191.

203. Hutchinson WL, Du MQ, Johnson PJ, Williams R. Fucosyltransferases: differential plasma and tissue alterations in hepatocellular carcinoma and cirrhosis. Hepatol Baltim Md. 1991 Apr;13(4):683–8.

204. Dennis JW, Laferté S, Waghorne C, Breitman ML, Kerbel RS. Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. Science. 1987 May 1;236(4801):582–5.

205. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW. Suppression of tumor growth and metastasis in Mgat5-deficient mice. Nat Med. 2000 Mar;6(3):306–12.

206. Nakagoe T, Fukushima K, Nanashima A, Sawai T, Tsuji T, Jibiki MA, et al. Comparison of the expression of ABH/Lewis-related antigens in polypoid and non-polypoid growth types of colorectal carcinoma. J Gastroenterol Hepatol. 2001 Feb;16(2):176–83.

207. Yuan M, Itzkowitz SH, Palekar A, Shamsuddin AM, Phelps PC, Trump BF, et al. Distribution of blood group antigens A, B, H, Lewisa, and Lewisb in human normal, fetal, and malignant colonic tissue. Cancer Res. 1985 Sep;45(9):4499–511.

208. Gwin JL, Klein-Szanto AJ, Zhang SY, Agarwal P, Rogatko A, Keller SM. Loss of blood group antigen A in non-small cell lung cancer. Ann Surg Oncol. 1994 Sep;1(5):423–7.

209. Miwa HE, Song Y, Alvarez R, Cummings RD, Stanley P. The bisecting GlcNAc in cell growth control and tumor progression. Glycoconj J. 2012 Dec;29(8-9):609–18.

210. Chrostek L, Cylwik B. [The alteration of proteins glycosylation in liver diseases]. Pol Merkur Lek Organ Pol Tow Lek. 2011 Jul;31(181):60–4.

211. Shirato K, Nakajima K, Korekane H, Takamatsu S, Gao C, Angata T, et al. Hypoxic regulation of glycosylation via the N-acetylglucosamine cycle. J Clin Biochem Nutr. 2011 Jan;48(1):20–5.

212. An G, Wei B, Xia B, McDaniel JM, Ju T, Cummings RD, et al. Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans. J Exp Med. 2007 Jun 11;204(6):1417–29.

213. Takada A, Ohmori K, Yoneda T, Tsuyuoka K, Hasegawa A, Kiso M, et al. Contribution of carbohydrate antigens sialyl Lewis A and sialyl Lewis X to adhesion of human cancer cells to vascular endothelium. Cancer Res. 1993 Jan 15;53(2):354–61.

214. Kim YJ, Borsig L, Varki NM, Varki A. P-selectin deficiency attenuates tumor growth and metastasis. Proc Natl Acad Sci U S A. 1998 Aug 4;95(16):9325–30.

215. Couldrey C, Green JE. Metastases: the glycan connection. Breast Cancer Res BCR. 2000;2(5):321–3.

216. Uemura T, Shiozaki K, Yamaguchi K, Miyazaki S, Satomi S, Kato K, et al. Contribution of sialidase NEU1 to suppression of metastasis of human colon cancer cells through desialylation of integrin beta4. Oncogene. 2009 Mar 5;28(9):1218–29.

217. Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. Cell surface protein glycosylation in cancer. Proteomics. 2014 Mar;14(4-5):525–46.

218. Stephens E, Maslen SL, Green LG, Williams DH. Fragmentation characteristics of neutral N-linked glycans using a MALDI-TOF/TOF tandem mass spectrometer. Anal Chem. 2004 Apr 15;76(8):2343–54.

219. Morelle W, Flahaut C, Michalski J-C, Louvet A, Mathurin P, Klein A. Mass spectrometric approach for screening modifications of total serum N-glycome in human diseases: application to cirrhosis. Glycobiology. 2006 Apr;16(4):281–93.

220. Morelle W, Faid V, Chirat F, Michalski J-C. Analysis of N- and O-linked glycans from glycoproteins using MALDI-TOF mass spectrometry. Methods Mol Biol Clifton NJ. 2009;534:5–21.

221. Frost DC, Li L. Recent advances in mass spectrometry-based glycoproteomics. Adv Protein Chem Struct Biol. 2014;95:71–123.

222. Dwek RA, Edge CJ, Harvey DJ, Wormald MR, Parekh RB. Analysis of glycoprotein-associated oligosaccharides. Annu Rev Biochem. 1993;62:65–100.

223. Zauner G, Kozak RP, Gardner RA, Fernandes DL, Deelder AM, Wuhrer M. Protein O-glycosylation analysis. Biol Chem [Internet]. 2012 Jan 1 [cited 2015 Feb 12];393(8).

Available from: http://www.degruyter.com/view/j/bchm.2012.393.issue-8/hsz-2012-0144/hsz-2012-0144.xml

224. Iyer RN, Carlson DM. Alkaline borohydride degradation of blood group H substance. Arch Biochem Biophys. 1971 Jan;142(1):101–5.

225. Maniatis S, Zhou H, Reinhold V. Rapid de-O-glycosylation concomitant with peptide labeling using microwave radiation and an alkyl amine base. Anal Chem. 2010 Mar 15;82(6):2421–5.

226. Huang Y, Mechref Y, Novotny MV. Microscale nonreductive release of O-linked glycans for subsequent analysis through MALDI mass spectrometry and capillary electrophoresis. Anal Chem. 2001 Dec 15;73(24):6063–9.

227. Yu G, Zhang Y, Zhang Z, Song L, Wang P, Chai W. Effect and limitation of excess ammonium on the release of O-glycans in reducing forms from glycoproteins under mild alkaline conditions for glycomic and functional analysis. Anal Chem. 2010 Nov 15;82(22):9534–42.

228. Miura Y, Kato K, Takegawa Y, Kurogochi M, Furukawa J-I, Shinohara Y, et al. Glycoblotting-assisted O-glycomics: ammonium carbamate allows for highly efficient o-glycan release from glycoproteins. Anal Chem. 2010 Dec 15;82(24):10021–9.

229. Merry AH, Neville DCA, Royle L, Matthews B, Harvey DJ, Dwek RA, et al. Recovery of intact 2-aminobenzamide-labeled O-glycans released from glycoproteins by hydrazinolysis. Anal Biochem. 2002 May 1;304(1):91–9.

230. Ruhaak LR, Zauner G, Huhn C, Bruggink C, Deelder AM, Wuhrer M. Glycan labeling strategies and their use in identification and quantification. Anal Bioanal Chem. 2010 Aug;397(8):3457–81.

231. Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB. Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. Anal Biochem. 1995 Sep 20;230(2):229–38.

232. Furukawa J, Fujitani N, Araki K, Takegawa Y, Kodama K, Shinohara Y. A versatile method for analysis of serine/threonine posttranslational modifications by β-elimination in the presence of pyrazolone analogues. Anal Chem. 2011 Dec 1;83(23):9060–7.

233. Zaia J. Mass spectrometry of oligosaccharides. Mass Spectrom Rev. 2004 Jun;23(3):161–227.

234. Goetz JA, Novotny MV, Mechref Y. Enzymatic/chemical release of O-glycans allowing MS analysis at high sensitivity. Anal Chem. 2009 Dec 1;81(23):9546–52.

235. Shah P, Yang S, Sun S, Aiyetan P, Yarema KJ, Zhang H. Mass spectrometric analysis of sialylated glycans with use of solid-phase labeling of sialic acids. Anal Chem. 2013 Apr 2;85(7):3606–13.

236. Harvey DJ, Baruah K, Scanlan CN. Application of negative ion MS/MS to the identification of N-glycans released from carcinoembryonic antigen cell adhesion molecule 1 (CEACAM1). J Mass Spectrom JMS. 2009 Jan;44(1):50–60.

237. Reiding KR, Blank D, Kuijper DM, Deelder AM, Wuhrer M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. Anal Chem. 2014 May 28;

238. Mechref Y, Kang P, Novotny MV. Solid-phase permethylation for glycomic analysis. Methods Mol Biol Clifton NJ. 2009;534:53–64.

239. Zaia J. Mass Spectrometry and Glycomics. OMICS J Integr Biol. 2010 Aug;14(4):401–18.

240. An HJ, Gip P, Kim J, Wu S, Park KW, McVaugh CT, et al. Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. Mol Cell Proteomics MCP. 2012 Apr;11(4):M111.010660.

241. Pabst M, Altmann F. Glycan analysis by modern instrumental methods. Proteomics. 2011 Feb;11(4):631–43.

242. Walker SH, Carlisle BC, Muddiman DC. Systematic comparison of reverse phase and hydrophilic interaction liquid chromatography platforms for the analysis of N-linked glycans. Anal Chem. 2012 Oct 2;84(19):8198–206.

243. Aich U, Birznieks I, Saba J, Liu X, Viner R, Hao Z, et al. Novel Glycan Column Technology for the LC-MS Analysis of Labeled and Native N-Glycans Released from Proteins and Antibodies. Thermo Fisher Scientific;

244. Stavenhagen K, Kolarich D, Wuhrer M. Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography–Mass Spectrometry. Chromatographia. 2015 Mar;78(5-6):307–20.

245. Dreisewerd K, Kölbl S, Peter-Katalinić J, Berkenkamp S, Pohlentz G. Analysis of native milk oligosaccharides directly from thin-layer chromatography plates by matrix-assisted laser desorption/ionization orthogonal-time-of-flight mass spectrometry with a glycerol matrix. J Am Soc Mass Spectrom. 2006 Feb;17(2):139–50.

246. Harvey DJ, Royle L, Radcliffe CM, Rudd PM, Dwek RA. Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. Anal Biochem. 2008 May 1;376(1):44–60.

247. Harvey DJ. Structural determination of N-linked glycans by matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry. Proteomics. 2005 May;5(7):1774–86.

248. Harvey DJ. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: An update covering the period 1999-2000. Mass Spectrom Rev. 2006 Aug;25(4):595–662.

249. Selman MHJ, McDonnell LA, Palmblad M, Ruhaak LR, Deelder AM, Wuhrer M. Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem. 2010 Feb 1;82(3):1073–81.

250. Selman MHJ, Hoffmann M, Zauner G, McDonnell LA, Balog CIA, Rapp E, et al. MALDI-TOF-MS analysis of sialylated glycans and glycopeptides using 4-chloro-α-cyanocinnamic acid matrix. Proteomics. 2012 May;12(9):1337–48.

251. Laštovičková M, Chmelik J, Bobalova J. The combination of simple MALDI matrices for the improvement of intact glycoproteins and glycans analysis. Int J Mass Spectrom. 2009 Mar;281(1-2):82–8.

252. Zhou S, Hu Y, DeSantos-Garcia JL, Mechref Y. Quantitation of Permethylated N-Glycans through Multiple-Reaction Monitoring (MRM) LC-MS/MS. J Am Soc Mass Spectrom. 2015 Apr;26(4):596–603.

253. Kronewitter SR, An HJ, de Leoz ML, Lebrilla CB, Miyamoto S, Leiserowitz GS. The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. Proteomics. 2009 Jun;9(11):2986–94.

254. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj J. 1988;5(4):397–409.

255. Harvey DJ, Bateman RH, Green MR. High-energy collision-induced fragmentation of complex oligosaccharides ionized by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom JMS. 1997 Feb;32(2):167–87.

256. Cancilla MT, Penn SG, Carroll JA, Lebrilla CB. Coordination of Alkali Metals to Oligosaccharides Dictates Fragmentation Behavior in Matrix Assisted Laser Desorption Ionization/Fourier Transform Mass Spectrometry. J Am Chem Soc. 1996 Jan;118(28):6736–45.

257. Thaysen-Andersen M, Packer NH. Advances in LC-MS/MS-based glycoproteomics: Getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. Biochim Biophys Acta. 2014 May 12;1844(9):1437–52.

258. Pan S, Chen R, Aebersold R, Brentnall TA. Mass spectrometry based glycoproteomics--from a proteomics perspective. Mol Cell Proteomics MCP. 2011 Jan;10(1):R110.003251.

259. Hua S, Hu CY, Kim BJ, Totten SM, Oh MJ, Yun N, et al. Glyco-analytical multispecific proteolysis (Glyco-AMP): a simple method for detailed and quantitative Glycoproteomic characterization. J Proteome Res. 2013 Oct 4;12(10):4414–23.

260. Dodds ED, Seipert RR, Clowers BH, German JB, Lebrilla CB. Analytical performance of immobilized pronase for glycopeptide footprinting and implications for surpassing reductionist glycoproteomics. J Proteome Res. 2009 Feb;8(2):502–12.

261. Sun Z, Sun D, Wang F, Cheng K, Zhang Z, Xu B, et al. Differential analysis of N-glycoproteome between hepatocellular carcinoma and normal human liver tissues by combination of multiple protease digestion and solid phase based labeling. Clin Proteomics. 2014;11(1):26.

262. Sun B, Ranish JA, Utleg AG, White JT, Yan X, Lin B, et al. Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. Mol Cell Proteomics MCP. 2007 Jan;6(1):141–9.

263. Gbormittah FO, Lee LY, Taylor K, Hancock WS, Iliopoulos O. Comparative Studies of the Proteome, Glycoproteome, and N-Glycome of Clear Cell Renal Cell Carcinoma Plasma before and after Curative Nephrectomy. J Proteome Res. 2014 Nov 7;13(11):4889–900.

264. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics MCP. 2002 Nov;1(11):845–67.

265. Wuhrer M, Catalina MI, Deelder AM, Hokke CH. Glycoproteomics based on tandem mass spectrometry of glycopeptides. J Chromatogr B. 2007 Apr;849(1-2):115–28.

266. Ongay S, Boichenko A, Govorukhina N, Bischoff R. Glycopeptide enrichment and separation for protein glycosylation analysis. J Sep Sci. 2012 Sep;35(18):2341–72.

267. Zhang Y, Jiao J, Yang P, Lu H. Mass spectrometry-based N-glycoproteomics for cancer biomarker discovery. Clin Proteomics. 2014;11(1):18.

268. Kubota K, Sato Y, Suzuki Y, Goto-Inoue N, Toda T, Suzuki M, et al. Analysis of glycopeptides using lectin affinity chromatography with MALDI-TOF mass spectrometry. Anal Chem. 2008 May 15;80(10):3693–8.

269. Drake RR, Schwegler EE, Malik G, Diaz J, Block T, Mehta A, et al. Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. Mol Cell Proteomics MCP. 2006 Oct;5(10):1957–67.

270. Tian Y, Zhang H. Characterization of disease-associated N-linked glycoproteins. Proteomics. 2013 Feb;13(3-4):504–11.

271. Kim EH, Misek DE. Glycoproteomics-Based Identification of Cancer Biomarkers. Int J Proteomics. 2011;2011:1–10.

272. Kaji H, Isobe T. Liquid Chromatography/Mass Spectrometry (LC/MS)-Based Glycoproteomics Technologies for Cancer Biomarker Discovery. Clin Proteomics. 2008 Jun;4(1-2):14–24.

273. Hirao Y, Ogasawara, S, Togayachi A. Identification of Core Proteins Carrying the Sialyl Lewis a Epitope in Pancreatic Cancers Identification of Core Proteins Carrying the Sialyl

Lewis a Epitope in Pancreatic Cancers. J Mol Biomark Diagn [Internet]. 2012 [cited 2015 Jan 28];03(03). Available from: http://www.omicsonline.org/2155-9929/2155-9929-3-124.digital/2155-9929-3-124.html

274. Cummings RD. The repertoire of glycan determinants in the human glycome. Mol Biosyst. 2009 Oct;5(10):1087–104.

275. Heimburg-Molinaro J, Priest JW, Live D, Boons G-J, Song X, Cummings RD, et al. Microarray analysis of the human antibody response to synthetic Cryptosporidium glycopeptides. Int J Parasitol. 2013 Oct;43(11):901–7.

276. Nilsson J, Larson G. Sialic Acid Capture-and-Release and LC-MSn Analysis of Glycopeptides. In: Kohler JJ, Patrie SM, editors. Mass Spectrometry of Glycoproteins [Internet]. Totowa, NJ: Humana Press; 2013 [cited 2014 Sep 17]. p. 79–100. Available from: http://link.springer.com/10.1007/978-1-62703-146-2_7

277. Zhang H, Li X, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol. 2003 Jun;21(6):660–6.

278. Petrescu A-J, Milac A-L, Petrescu SM, Dwek RA, Wormald MR. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. Glycobiology. 2004 Feb;14(2):103–14.

279. Ramachandran P, Boontheung P, Pang E, Yan W, Wong DT, Loo JA. Comparison of N-linked Glycoproteins in Human Whole Saliva, Parotid, Submandibular, and Sublingual Glandular Secretions Identified using Hydrazide Chemistry and Mass Spectrometry. Clin Proteomics. 2008 Dec;4(3-4):80–104.

280. Lee MCG, Sun B. Glycopeptide capture for cell surface proteomics. J Vis Exp JoVE. 2014;(87).

281. Li QK, Shah P, Li Y, Aiyetan PO, Chen J, Yung R, et al. Glycoproteomic analysis of bronchoalveolar lavage (BAL) fluid identifies tumor-associated glycoproteins from lung adenocarcinoma. J Proteome Res. 2013 Aug 2;12(8):3689–96.

282. Song E, Zhu R, Hammoud ZT, Mechref Y. LC-MS/MS quantitation of esophagus disease blood serum glycoproteins by enrichment with hydrazide chemistry and lectin affinity chromatography. J Proteome Res. 2014 Nov 7;13(11):4808–20.

283. Pan S, Chen R, Tamura Y, Crispin DA, Lai LA, May DH, et al. Quantitative Glycoproteomics Analysis Reveals Changes in N-Glycosylation Level Associated with Pancreatic Ductal Adenocarcinoma. J Proteome Res. 2014 Jan 28;

284. Wright HT. Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. Crit Rev Biochem Mol Biol. 1991;26(1):1–52.

285. Palmisano G, Melo-Braga MN, Engholm-Keller K, Parker BL, Larsen MR. Chemical Deamidation: A Common Pitfall in Large-Scale N-Linked Glycoproteomic Mass Spectrometry-Based Analyses. J Proteome Res. 2012 Mar 2;11(3):1949–57.

286. Nilsson J, Rüetschi U, Halim A, Hesse C, Carlsohn E, Brinkmalm G, et al. Enrichment of glycopeptides for glycan structure and attachment site identification. Nat Methods. 2009 Nov;6(11):809–11.

287. Sparbier K, Koch S, Kessler I, Wenzel T, Kostrzewa M. Selective isolation of glycoproteins and glycopeptides for MALDI-TOF MS detection supported by magnetic particles. J Biomol Tech JBT. 2005 Dec;16(4):407–13.

288. Sletten EM, Bertozzi CR. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew Chem Int Ed Engl. 2009;48(38):6974–98.

289. Zhang X, Zhang Y. Applications of azide-based bioorthogonal click chemistry in glycobiology. Mol Basel Switz. 2013;18(6):7145–59.

290. Laughlin ST, Agard NJ, Baskin JM, Carrico IS, Chang PV, Ganguli AS, et al. Metabolic labeling of glycans with azido sugars for visualization and glycoproteomics. Methods Enzymol. 2006;415:230–50.

291. Jacobs CL, Yarema KJ, Mahal LK, Nauman DA, Charters NW, Bertozzi CR. Metabolic labeling of glycoproteins with chemical tags through unnatural sialic acid biosynthesis. Methods Enzymol. 2000;327:260–75.

292. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. In Vivo Imaging of Membrane-Associated Glycans in Developing Zebrafish. Science. 2008 May 2;320(5876):664–7.

293. Yang L, Nyalwidhe JO, Guo S, Drake RR, Semmes OJ. Targeted identification of metastasis-associated cell-surface sialoglycoproteins in prostate cancer. Mol Cell Proteomics MCP. 2011 Jun;10(6):M110.007294.

294. Click-iT Protein Enrichment Kit. Invitrogen; 2011.

295. Roper SM, Zemskova M, Neely BA, Martin A, Gao P, Jones EE, et al. Targeted glycoprotein enrichment and identification in stromal cell secretomes using azido sugar metabolic labeling. Proteomics Clin Appl. 2013 Jun;7(5-6):367–71.

296. Saxon E. Cell Surface Engineering by a Modified Staudinger Reaction. Science. 2000 Mar 17;287(5460):2007–10.

297. Hang HC, Yu C, Kato DL, Bertozzi CR. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. Proc Natl Acad Sci U S A. 2003 Dec 9;100(25):14846–51.

298. Vocadlo DJ, Hang HC, Kim E-J, Hanover JA, Bertozzi CR. A chemical approach for identifying O-GlcNAc-modified proteins in cells. Proc Natl Acad Sci U S A. 2003 Aug 5;100(16):9116–21.

299. Prescher JA, Dube DH, Bertozzi CR. Chemical remodelling of cell surfaces in living animals. Nature. 2004 Aug 19;430(7002):873–7.

300. Dube DH, Prescher JA, Quang CN, Bertozzi CR. Probing mucin-type O-linked glycosylation in living animals. Proc Natl Acad Sci U S A. 2006 Mar 28;103(13):4819–24.

301. Luchansky SJ, Argade S, Hayes BK, Bertozzi CR. Metabolic functionalization of recombinant glycoproteins. Biochemistry (Mosc). 2004 Sep 28;43(38):12358–66.

302. Rabuka D, Hubbard SC, Laughlin ST, Argade SP, Bertozzi CR. A chemical reporter strategy to probe glycoprotein fucosylation. J Am Chem Soc. 2006 Sep 20;128(37):12078–9.

303. Wohlgemuth J, Karas M, Jiang W, Hendriks R, Andrecht S. Enhanced glyco-profiling by specific glycopeptide enrichment and complementary monolithic nano-LC (ZIC-HILIC/RP18e)/ESI-MS analysis. J Sep Sci. 2010 Mar;33(6-7):880–90.

304. Parker BL, Thaysen-Andersen M, Solis N, Scott NE, Larsen MR, Graham ME, et al. Site-Specific Glycan-Peptide Analysis for Determination of *N* -Glycoproteome Heterogeneity. J Proteome Res. 2013 Dec 6;12(12):5791–800.

305. Snovida SI, Bodnar ED, Viner R, Saba J, Perreault H. A simple cellulose column procedure for selective enrichment of glycopeptides and characterization by nano LC coupled with electron-transfer and high-energy collisional-dissociation tandem mass spectrometry. Carbohydr Res. 2010 Apr 19;345(6):792–801.

306. Mysling S, Palmisano G, Højrup P, Thaysen-Andersen M. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. Anal Chem. 2010 Jul 1;82(13):5598–609.

307. Martin GS. Cell signaling and cancer. Cancer Cell. 2003 Sep;4(3):167–74.

308. Dedvisitsakul P, Jacobsen S, Svensson B, Bunkenborg J, Finnie C, Hägglund P. Glycopeptide Enrichment Using a Combination of ZIC-HILIC and Cotton Wool for Exploring the Glycoproteome of Wheat Flour Albumins. J Proteome Res. 2014 Apr;140401073008005.

309. Chen R, Seebun D, Ye M, Zou H, Figeys D. Site-specific Characterization of Cell Membrane N-Glycosylation with Integrated Hydrophilic Interaction Chromatography Solid Phase Extraction and LC-MS/MS. J Proteomics. 2014 Apr 7;

310. Zhu J, Sun Z, Cheng K, Chen R, Ye M, Xu B, et al. Comprehensive mapping of protein N-glycosylation in human liver by combining hydrophilic interaction chromatography and hydrazide chemistry. J Proteome Res. 2014 Mar 7;13(3):1713–21.

311. Larsen MR, Jensen SS, Jakobsen LA, Heegaard NHH. Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. Mol Cell Proteomics MCP. 2007 Oct;6(10):1778–87.

312. Takakura D, Harazono A, Hashii N, Kawasaki N. Selective glycopeptide profiling by acetone enrichment and LC/MS. J Proteomics. 2014 Feb 12;

313. Alvarez-Manilla G, Atwood J, Guo Y, Warren NL, Orlando R, Pierce M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. J Proteome Res. 2006 Mar;5(3):701–8.

314. Li X, Jiang J, Zhao X, Wang J, Han H, Zhao Y, et al. N-glycoproteome analysis of the secretome of human metastatic hepatocellular carcinoma cell lines combining hydrazide chemistry, HILIC enrichment and mass spectrometry. PloS One. 2013;8(12):e81921.

315. Morelle W, Michalski J-C. The Mass Spectrometric Analysis of Glycoproteins and their Glycan Structures. Curr Anal Chem. 2005 Jan 1;1(1):29–57.

316. Thaysen-Andersen M, Mysling S, Højrup P. Site-specific glycoprofiling of N-linked glycopeptides using MALDI-TOF MS: strong correlation between signal strength and glycoform quantities. Anal Chem. 2009 May 15;81(10):3933–43.

317. Irungu J, Go EP, Zhang Y, Dalpathado DS, Liao H-X, Haynes BF, et al. Comparison of HPLC/ESI-FTICR MS versus MALDI-TOF/TOF MS for glycopeptide analysis of a highly glycosylated HIV envelope glycoprotein. J Am Soc Mass Spectrom. 2008 Aug;19(8):1209–20.

318. Kontro H, Joenväärä S, Haglund C, Renkonen R. Comparison of sialylated N-glycopeptide levels in serum of pancreatic cancer patients, acute pancreatitis patients and healthy controls. PROTEOMICS. 2014 May;n/a – n/a.

319. Ye H, Boyne MT, Buhse LF, Hill J. Direct approach for qualitative and quantitative characterization of glycoproteins using tandem mass tags and an LTQ Orbitrap XL electron transfer dissociation hybrid mass spectrometer. Anal Chem. 2013 Feb 5;85(3):1531–9.

320. Saba J, Dutta S, Hemenway E, Viner R. Increasing the Productivity of Glycopeptides Analysis by Using Higher-Energy Collision Dissociation-Accurate Mass-Product-Dependent Electron Transfer Dissociation. Int J Proteomics. 2012;2012:1–7.

321. Yin X, Bern M, Xing Q, Ho J, Viner R, Mayr M. Glycoproteomic analysis of the secretome of human endothelial cells. Mol Cell Proteomics MCP. 2013 Apr;12(4):956–78.

322. Wu S-W, Pu T-H, Viner R, Khoo K-H. Novel LC-MS^2 product dependent parallel data acquisition function and data analysis workflow for sequencing and identification of intact glycopeptides. Anal Chem. 2014 Jun 3;86(11):5478–86.

323. Gosho M, Nagashima K, Sato Y. Study designs and statistical analyses for biomarker research. Sensors. 2012;12(7):8966–86.

324. Füzéry AK, Levin J, Chan MM, Chan DW. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. Clin Proteomics. 2013;10(1):13.

325. Kuzmanov U, Kosanam H, Diamandis EP. The sweet and sour of serological glycoprotein tumor biomarker quantification. BMC Med. 2013;11:31.

326. Malati T. Tumour markers: An overview. Indian J Clin Biochem. 2007 Sep;22(2):17–31.

327. Narimatsu H. Strategy for development of clinically useful glyco-biomarkers. Glycoconj J. 2014 Oct;31(6-7):403–7.

328. Tian Y, Gurley K, Meany DL, Kemp CJ, Zhang H. N-linked glycoproteomic analysis of formalin-fixed and paraffin-embedded tissues. J Proteome Res. 2009 Apr;8(4):1657–62.

329. Barry MJ. Screening for Prostate Cancer — The Controversy That Refuses to Die. N Engl J Med. 2009 Mar 26;360(13):1351–4.

330. Cooperberg MR, Broering JM, Kantoff PW, Carroll PR. Contemporary trends in low risk prostate cancer: risk assessment and treatment. J Urol. 2007 Sep;178(3 Pt 2):S14–9.

331. Hirano K, Matsuda A, Shirai T, Furukawa K. Expression of LacdiNAc groups on N-glycans among human tumors is complex. BioMed Res Int. 2014;2014:981627.

332. Koprowski H, Steplewski Z, Mitchell K, Herlyn M, Herlyn D, Fuhrer P. Colorectal carcinoma antigens detected by hybridoma antibodies. Somatic Cell Genet. 1979 Nov;5(6):957–71.

333. Fong ZV, Winter JM. Biomarkers in pancreatic cancer: diagnostic, prognostic, and predictive. Cancer J Sudbury Mass. 2012 Dec;18(6):530–8.

334. Magnani JL, Nilsson B, Brockhaus M, Zopf D, Steplewski Z, Koprowski H, et al. A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-N-fucopentaose II. J Biol Chem. 1982 Dec 10;257(23):14365–9.

335. Ho JJ, Siddiki B, Kim YS. Association of sialyl-Lewis(a) and sialyl-Lewis(x) with MUC-1 apomucin ina pancreatic cancer cell line. Cancer Res. 1995 Aug 15;55(16):3659–63.

336. Pérez-Garay M, Arteta B, Llop E, Cobler L, Pagès L, Ortiz R, et al. α2,3-Sialyltransferase ST3Gal IV promotes migration and metastasis in pancreatic adenocarcinoma cells and tends to be highly expressed in pancreatic adenocarcinoma tissues. Int J Biochem Cell Biol. 2013 Aug;45(8):1748–57.

337. Ugorski M, Laskowska A. Sialyl Lewis(a): a tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells. Acta Biochim Pol. 2002;49(2):303–11.

338. Iwai K, Ishikura H, Kaji M, Sugiura H, Ishizu A, Takahashi C, et al. Importance of E-selectin (ELAM-1) and sialyl Lewis(a) in the adhesion of pancreatic carcinoma cells to activated endothelium. Int J Cancer J Int Cancer. 1993 Jul 30;54(6):972–7.

339. Sawada T, Ho JJ, Chung YS, Sowa M, Kim YS. E-selectin binding by pancreatic tumor cells is inhibited by cancer sera. Int J Cancer J Int Cancer. 1994 Jun 15;57(6):901–7.

340. Kłopocki AG, Laskowska A, Antoniewicz-Papis J, Duk M, Lisowska E, Ugorski M. Role of sialosyl Lewis(a) in adhesion of colon cancer cells--the antisense RNA approach. Eur J Biochem FEBS. 1998 Apr 1;253(1):309–18.

341. Marrelli D, Caruso S, Pedrazzani C, Neri A, Fernandes E, Marini M, et al. CA19-9 serum levels in obstructive jaundice: clinical value in benign and malignant conditions. Am J Surg. 2009 Sep;198(3):333–9.

342. Parra JL, Kaplan S, Barkin JS. Elevated CA 19-9 caused by Hashimoto's thyroiditis: review of the benign causes of increased CA 19-9 level. Dig Dis Sci. 2005 Apr;50(4):694–5.

343. Nie S, Lo A, Wu J, Zhu J, Tan Z, Simeone DM, et al. Glycoprotein biomarker panel for pancreatic cancer discovered by quantitative proteomics analysis. J Proteome Res. 2014 Apr 4;13(4):1873–84.

344. Kim J-E, Lee KT, Lee JK, Paik SW, Rhee JC, Choi KW. Clinical usefulness of carbohydrate antigen 19-9 as a screening test for pancreatic cancer in an asymptomatic population. J Gastroenterol Hepatol. 2004 Feb;19(2):182–6.

345. Chang C-Y, Huang S-P, Chiu H-M, Lee Y-C, Chen M-F, Lin J-T. Low efficacy of serum levels of CA 19-9 in prediction of malignant diseases in asymptomatic population in Taiwan. Hepatogastroenterology. 2006 Feb;53(67):1–4.

346. Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. J Gastrointest Oncol. 2012 Jun;3(2):105–19.

347. Schlieman MG, Ho HS, Bold RJ. Utility of tumor markers in determining resectability of pancreatic cancer. Arch Surg Chic Ill 1960. 2003 Sep;138(9):951–5; discussion 955–6.

348. Waraya M, Yamashita K, Katagiri H, Ishii K, Takahashi Y, Furuta K, et al. Preoperative serum CA19-9 and dissected peripancreatic tissue margin as determiners of long-term survival in pancreatic cancer. Ann Surg Oncol. 2009 May;16(5):1231–40.

349. Rich N, Singal AG. Hepatocellular carcinoma tumour markers: current role and expectations. Best Pract Res Clin Gastroenterol. 2014 Oct;28(5):843–53.

350. Aoyagi Y, Suzuki Y, Isemura M, Nomoto M, Sekine C, Igarashi K, et al. The fucosylation index of alpha-fetoprotein and its usefulness in the early diagnosis of hepatocellular carcinoma. Cancer. 1988 Feb 15;61(4):769–74.

351. Yamashita K, Taketa K, Nishi S, Fukushima K, Ohkura T. Sugar chains of human cord serum alpha-fetoprotein: characteristics of N-linked sugar chains of glycoproteins produced in human liver and hepatocellular carcinomas. Cancer Res. 1993 Jul 1;53(13):2970–5.

352. Aoyagi Y, Suzuki Y, Igarashi K, Saitoh A, Oguro M, Yokota T, et al. The usefulness of simultaneous determinations of glucosaminylation and fucosylation indices of alpha-fetoprotein in the differential diagnosis of neoplastic diseases of the liver. Cancer. 1991 May 1;67(9):2390–4.

353. Guillaumond F, Leca J, Olivares O, Lavaut M-N, Vidal N, Berthezène P, et al. Strengthened glycolysis under hypoxia supports tumor symbiosis and hexosamine biosynthesis in pancreatic adenocarcinoma. Proc Natl Acad Sci U S A. 2013 Mar 5;110(10):3919–24.

354. Upadhyay M, Samal J, Kandpal M, Singh OV, Vivekanandan P. The Warburg effect: insights from the past decade. Pharmacol Ther. 2013 Mar;137(3):318–30.

355. Guillaumond F, Iovanna JL, Vasseur S. Pancreatic tumor cell metabolism: Focus on glycolysis and its connected metabolic pathways. Arch Biochem Biophys. 2014 Mar;545:69–73.

356. Kannagi R, Sakuma K, Miyazaki K, Lim K-T, Yusa A, Yin J, et al. Altered expression of glycan genes in cancers induced by epigenetic silencing and tumor hypoxia: clues in the ongoing search for new tumor markers. Cancer Sci. 2010 Mar;101(3):586–93.

357. Koike T, Kimura N, Miyazaki K, Yabuta T, Kumamoto K, Takenoshita S, et al. Hypoxia induces adhesion molecules on cancer cells: A missing link between Warburg effect and induction of selectin-ligand carbohydrates. Proc Natl Acad Sci U S A. 2004 May 25;101(21):8132–7.

358. Yin J, Hashimoto A, Izawa M, Miyazaki K, Chen G-Y, Takematsu H, et al. Hypoxic culture induces expression of sialin, a sialic acid transporter, and cancer-associated gangliosides containing non-human sialic acid on human cancer cells. Cancer Res. 2006 Mar 15;66(6):2937–45.

359. Graeber TG, Osmanian C, Jacks T, Housman DE, Koch CJ, Lowe SW, et al. Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumours. Nature. 1996 Jan 4;379(6560):88–91.

360. Yun J, Rago C, Cheong I, Pagliarini R, Angenendt P, Rajagopalan H, et al. Glucose deprivation contributes to the development of KRAS pathway mutations in tumor cells. Science. 2009 Sep 18;325(5947):1555–9.

361. Ying H, Kimmelman AC, Lyssiotis CA, Hua S, Chu GC, Fletcher-Sananikone E, et al. Oncogenic Kras Maintains Pancreatic Tumors through Regulation of Anabolic Glucose Metabolism. Cell. 2012 Apr;149(3):656–70.

362. Pérez-Garay M, Arteta B, Pagès L, de Llorens R, de Bolòs C, Vidal-Vanaclocha F, et al. alpha2,3-sialyltransferase ST3Gal III modulates pancreatic cancer cell motility and adhesion in vitro and enhances its metastatic potential in vivo. PloS One. 2010;5(9).

363. Bassagañas S, Carvalho S, Dias AM, Pérez-Garay M, Ortiz MR, Figueras J, et al. Pancreatic Cancer Cell Glycosylation Regulates Cell Adhesion and Invasion through the Modulation of α2β1 Integrin and E-Cadherin Function. St-Pierre Y, editor. PLoS ONE. 2014 May 30;9(5):e98595.

364. Pan S, Brentnall TA, Kelly K, Chen R. Tissue proteomics in pancreatic cancer study: discovery, emerging technologies, and challenges. Proteomics. 2013 Feb;13(3-4):710–21.

365. Chen R, Yi EC, Donohoe S, Pan S, Eng J, Cooke K, et al. Pancreatic Cancer Proteome: The Proteins That Underlie Invasion, Metastasis, and Immunologic Escape. Gastroenterology. 2005 Oct;129(4):1187–97.

366. Lu Z, Hu L, Evers S, Chen J, Shen Y. Differential expression profiling of human pancreatic adenocarcinoma and healthy pancreatic tissue. Proteomics. 2004 Dec;4(12):3975–88.

367. Bausch D, Thomas S, Mino-Kenudson M, Fernández-del CC, Bauer TW, Williams M, et al. Plectin-1 as a novel biomarker for pancreatic cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2011 Jan 15;17(2):302–9.

368. Pan S, Chen R, Reimel BA, Crispin DA, Mirzaei H, Cooke K, et al. Quantitative proteomics investigation of pancreatic intraepithelial neoplasia. ELECTROPHORESIS. 2009 Apr;30(7):1132–44.

369. Shen J, Person MD, Zhu J, Abbruzzese JL, Li D. Protein expression profiles in pancreatic adenocarcinoma compared with normal pancreatic tissue and tissue affected by pancreatitis as detected by two-dimensional gel electrophoresis and mass spectrometry. Cancer Res. 2004 Dec 15;64(24):9018–26.

370. Pan S, Chen R, Brand RE, Hawley S, Tamura Y, Gafken PR, et al. Multiplex targeted proteomic assay for biomarker detection in plasma: a pancreatic cancer biomarker case study. J Proteome Res. 2012 Mar 2;11(3):1937–48.

371. Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, et al. EMT and dissemination precede pancreatic tumor formation. Cell. 2012 Jan 20;148(1-2):349–61.

372. Shekouh AR, Thompson CC, Prime W, Campbell F, Hamlett J, Herrington CS, et al. Application of laser capture microdissection combined with two-dimensional electrophoresis for the discovery of differentially regulated proteins in pancreatic ductal adenocarcinoma. Proteomics. 2003 Oct;3(10):1988–2001.

373. Yue T, Goldstein IJ, Hollingsworth MA, Kaul K, Brand RE, Haab BB. The prevalence and nature of glycan alterations on specific proteins in pancreatic cancer patients revealed using antibody-lectin sandwich arrays. Mol Cell Proteomics MCP. 2009 Jul;8(7):1697–707.

374. Swanson BJ, McDermott KM, Singh PK, Eggers JP, Crocker PR, Hollingsworth MA. MUC1 is a counter-receptor for myelin-associated glycoprotein (Siglec-4a) and their interaction contributes to adhesion in pancreatic cancer perineural invasion. Cancer Res. 2007 Nov 1;67(21):10222–9.

375. Chaturvedi P, Singh AP, Chakraborty S, Chauhan SC, Bafna S, Meza JL, et al. MUC4 mucin interacts with and stabilizes the HER2 oncoprotein in human pancreatic cancer cells. Cancer Res. 2008 Apr 1;68(7):2065–70.

376. Grønborg M, Kristiansen TZ, Iwahori A, Chang R, Reddy R, Sato N, et al. Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. Mol Cell Proteomics MCP. 2006 Jan;5(1):157–71.

377. Pan S, Tamura Y, Chen R, May D, McIntosh MW, Brentnall TA. Large-scale quantitative glycoproteomics analysis of site-specific glycosylation occupancy. Mol Biosyst. 2012;8(11):2850.

378. Nakata D. Increased N-glycosylation of Asn^{88} in serum pancreatic ribonuclease 1 is a novel diagnostic marker for pancreatic cancer. Sci Rep. 2014;4:6715.

379. Chen R, Brentnall TA, Pan S, Cooke K, Moyes KW, Lane Z, et al. Quantitative proteomics analysis reveals that proteins differentially expressed in chronic pancreatitis are also frequently involved in pancreatic cancer. Mol Cell Proteomics MCP. 2007 Aug;6(8):1331–42.

380. Pan S, Chen R, Stevens T, Bronner MP, May D, Tamura Y, et al. Proteomics portrait of archival lesions of chronic pancreatitis. PloS One. 2011;6(11):e27574.

381. Porterfield M, Zhao P, Han H, Cunningham J, Aoki K, Von Hoff DD, et al. Discrimination between Adenocarcinoma and Normal Pancreatic Ductal Fluid by Proteomic and Glycomic Analysis. J Proteome Res. 2014 Feb 7;13(2):395–407.

382. Felix K, Hauck O, Fritz S, Hinz U, Schnölzer M, Kempf T, et al. Serum protein signatures differentiating autoimmune pancreatitis versus pancreatic cancer. PloS One. 2013;8(12):e82755.

383. Brand RE, Nolen BM, Zeh HJ, Allen PJ, Eloubeidi MA, Goldberg M, et al. Serum biomarker panels for the detection of pancreatic cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2011 Feb 15;17(4):805–16.

384. Li C, Zolotarevsky E, Thompson I, Anderson MA, Simeone DM, Casper JM, et al. A multiplexed bead assay for profiling glycosylation patterns on serum protein biomarkers of pancreatic cancer. Electrophoresis. 2011 Aug;32(15):2028–35.

385. Nouso K, Amano M, Ito YM, Miyahara K, Morimoto Y, Kato H, et al. Clinical utility of high-throughput glycome analysis in patients with pancreatic cancer. J Gastroenterol. 2013 Oct;48(10):1171–9.

386. National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. Transforming Glycoscience: A Roadmap for the Future [Internet]. Washington (DC): National Academies Press (US); 2012 [cited 2014 Apr 16]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK109958/

387. Magdeldin S, Yamamoto T. Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. Proteomics. 2012 Apr;12(7):1045–58.

388. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. Anal Chem. 1997 Dec 1;69(23):4751–60.

389. Seeley EH, Schwamborn K, Caprioli RM. Imaging of intact tissue sections: moving beyond the microscope. J Biol Chem. 2011 Jul 22;286(29):25459–66.

390. Römpp A, Spengler B. Mass spectrometry imaging with high resolution in mass and space. Histochem Cell Biol. 2013 Jun;139(6):759–83.

391. Seeley EH, Caprioli RM. 3D imaging by mass spectrometry: a new frontier. Anal Chem. 2012 Mar 6;84(5):2105–10.

392. Thiele H, Heldmann S, Trede D, Strehlow J, Wirtz S, Dreher W, et al. 2D and 3D MALDI-imaging: Conceptual strategies for visualization and data mining. Biochim Biophys Acta BBA - Proteins Proteomics. 2014 Jan;1844(1):117–37.

393. Schwamborn K, Caprioli RM. Molecular imaging by mass spectrometry — looking beyond classical histology. Nat Rev Cancer. 2010 Aug 19;10(9):639–46.

394. Balluff B, Schöne C, Höfler H, Walch A. MALDI imaging mass spectrometry for direct tissue analysis: technological advancements and recent applications. Histochem Cell Biol. 2011 Jul 30;136(3):227–44.

395. Cazares LH, Troyer D, Mendrinos S, Lance RA, Nyalwidhe JO, Beydoun HA, et al. Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 2 discriminates cancer from uninvolved prostate tissue. Clin Cancer Res Off J Am Assoc Cancer Res. 2009 Sep 1;15(17):5541–51.

396. Chaurand P, Cornett DS, Angel PM, Caprioli RM. From whole-body sections down to cellular level, multiscale imaging of phospholipids by MALDI mass spectrometry. Mol Cell Proteomics MCP. 2011 Feb;10(2):O110.004259.

397. Berry KAZ, Hankin JA, Barkley RM, Spraggins JM, Caprioli RM, Murphy RC. MALDI imaging of lipid biochemistry in tissues by mass spectrometry. Chem Rev. 2011 Oct 12;111(10):6491–512.

398. Castellino S, Groseclose MR, Wagner D. MALDI imaging mass spectrometry: bridging biology and chemistry in drug development. Bioanalysis. 2011 Nov;3(21):2427–41.

399. Nilsson A, Fehniger TE, Gustavsson L, Andersson M, Kenne K, Marko-Varga G, et al. Fine mapping the spatial distribution and concentration of unlabeled drugs within tissue micro-compartments using imaging mass spectrometry. PloS One. 2010;5(7):e11411.

400. Dennis JW, Nabi IR, Demetriou M. Metabolism, cell surface organization, and disease. Cell. 2009 Dec 24;139(7):1229–41.

401. Leymarie N, Zaia J. Effective use of mass spectrometry for glycan and glycopeptide structural analysis. Anal Chem. 2012 Apr 3;84(7):3040–8.

402. Haslam SM, North SJ, Dell A. Mass spectrometric analysis of N- and O-glycosylation of tissues and cells. Curr Opin Struct Biol. 2006 Oct;16(5):584–91.

403. Wada Y, Azadi P, Costello CE, Dell A, Dwek RA, Geyer H, et al. Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. Glycobiology. 2007 Apr;17(4):411–22.

404. Clemis EJ, Smith DS, Camenzind AG, Danell RM, Parker CE, Borchers CH. Quantitation of spatially-localized proteins in tissue samples using MALDI-MRM imaging. Anal Chem. 2012 Apr 17;84(8):3514–22.

405. Mehta A, Carrouée S, Conyers B, Jordan R, Butters T, Dwek RA, et al. Inhibition of hepatitis B virus DNA replication by imino sugars without the inhibition of the DNA polymerase: therapeutic implications. Hepatol Baltim Md. 2001 Jun;33(6):1488–95.

406. Guile GR, Rudd PM, Wing DR, Prime SB, Dwek RA. A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. Anal Biochem. 1996 Sep 5;240(2):210–26.

407. Glycan Structures Database [Internet]. Consortium for Functional Glycomics. 2010. Available from: http://www.functionalglycomics.org/fg/

408. Bern M, Brito AE, Pang P-C, Rekhi A, Dell A, Haslam SM. Polylactosaminoglycan glycomics: enhancing the detection of high-molecular-weight N-glycans in matrix-assisted laser desorption ionization time-of-flight profiles by matched filtering. Mol Cell Proteomics MCP. 2013 Apr;12(4):996–1004.

409. Lagarrigue M, Alexandrov T, Dieuset G, Perrin A, Lavigne R, Baulac S, et al. New analysis workflow for MALDI imaging mass spectrometry: application to the discovery and identification of potential markers of childhood absence epilepsy. J Proteome Res. 2012 Nov 2;11(11):5453–63.

410. Källback P, Shariatgorji M, Nilsson A, Andrén PE. Novel mass spectrometry imaging software assisting labeled normalization and quantitation of drugs and neuropeptides directly in tissue sections. J Proteomics. 2012 Aug 30;75(16):4941–51.

411. Jones EA, Shyti R, van Zeijl RJM, van Heiningen SH, Ferrari MD, Deelder AM, et al. Imaging mass spectrometry to visualize biomolecule distributions in mouse brain tissue following hemispheric cortical spreading depression. J Proteomics. 2012 Aug 30;75(16):5027–35.

412. Mechref Y, Hu Y, Garcia A, Zhou S, Desantos-Garcia JL, Hussein A. Defining putative glycan cancer biomarkers by MS. Bioanalysis. 2012 Oct;4(20):2457–69.

413. Wuhrer M. Glycomics using mass spectrometry. Glycoconj J. 2013 Jan;30(1):11–22.

414. Lemaire R, Desmons A, Tabet JC, Day R, Salzet M, Fournier I. Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. J Proteome Res. 2007 Apr;6(4):1295–305.

415. Schultz MJ, Swindall AF, Bellis SL. Regulation of the metastatic cell phenotype by sialylated glycans. Cancer Metastasis Rev. 2012 Dec;31(3-4):501–18.

416. Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer. 2005 Nov;5(11):845–56.

417. Cazares LH, Troyer DA, Wang B, Drake RR, Semmes OJ. MALDI tissue imaging: from biomarker discovery to clinical applications. Anal Bioanal Chem. 2011 Jul;401(1):17–27.

418. Dai C, Cazares LH, Wang L, Chu Y, Wang SL, Troyer DA, et al. Using boronolectin in MALDI-MS imaging for the histological analysis of cancer tissue expressing the sialyl Lewis X antigen. Chem Commun Camb Engl. 2011 Oct 7;47(37):10338–40.

419. Thompson SM, Craven RA, Nirmalan NJ, Harnden P, Selby PJ, Banks RE. Impact of pre-analytical factors on the proteomic analysis of formalin-fixed paraffin-embedded tissue. Proteomics Clin Appl. 2013 Apr;7(3-4):241–51.

420. Craven RA, Cairns DA, Zougman A, Harnden P, Selby PJ, Banks RE. Proteomic analysis of formalin-fixed paraffin-embedded renal tissue samples by label-free MS: assessment of overall technical variability and the impact of block age. Proteomics Clin Appl. 2013 Apr;7(3-4):273–82.

421. Wiśniewski JR, Duś K, Mann M. Proteomic workflow for analysis of archival formalin-fixed and paraffin-embedded clinical samples to a depth of 10 000 proteins. Proteomics Clin Appl. 2013 Apr;7(3-4):225–33.

422. Takikita M, Chung J-Y, Hewitt SM. Tissue microarrays enabling high-throughput molecular pathology. Curr Opin Biotechnol. 2007 Aug;18(4):318–25.

423. Franco R, Caraglia M, Facchini G, Abbruzzese A, Botti G. The role of tissue microarray in the era of target-based agents. Expert Rev Anticancer Ther. 2011 Jun;11(6):859–69.

424. Hewitt SM. The Application of Tissue Microarrays in the Validation of Microarray Results. Methods in Enzymology [Internet]. Elsevier; 2006 [cited 2015 Mar 10]. p. 400–15. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0076687906100208

425. Camp RL, Neumeister V, Rimm DL. A Decade of Tissue Microarrays: Progress in the Discovery and Validation of Cancer Biomarkers. J Clin Oncol. 2008 Oct 20;26(34):5630–7.

426. Van Cruijsen H, Gallegos Ruiz M, van der Valk P, de Gruijl TD, Giaccone G. Tissue micro array analysis of ganglioside N-glycolyl GM3 expression and signal transducer and activator of transcription (STAT)-3 activation in relation to dendritic cell infiltration and microvessel density in non-small cell lung cancer. BMC Cancer. 2009;9(1):180.

427. Chen C-Y, Jan Y-H, Juan Y-H, Yang C-J, Huang M-S, Yu C-J, et al. Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer. Proc Natl Acad Sci. 2013 Jan 8;110(2):630–5.

428. Kobayashi M, Nakayama J. Immunohistochemical Analysis of Carbohydrate Antigens in Chronic Inflammatory Gastrointestinal Diseases. Methods in Enzymology [Internet]. Elsevier; 2010 [cited 2015 Mar 10]. p. 271–89. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0076687910790169

429. Groseclose MR, Massion PP, Chaurand P, Caprioli RM. High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. Proteomics. 2008 Sep;8(18):3715–24.

430. Quaas A, Bahar AS, von Loga K, Seddiqi AS, Singer JM, Omidi M, et al. MALDI imaging on large-scale tissue microarrays identifies molecular features associated with tumour phenotype in oesophageal cancer. Histopathology. 2013 Jul

431. Powers TW, Jones EE, Betesh LR, Romano PR, Gao P, Copland JA, 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 Oct 15;85(20):9799–806.

432. Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans †. J Proteome Res. 2008 Apr;7(4):1650–9.

433. Shi C, Washington MK, Chaturvedi R, Drosos Y, Revetta FL, Weaver CJ, et al. Fibrogenesis in pancreatic cancer is a dynamic process regulated by macrophage–stellate cell interaction. Lab Invest. 2014 Apr;94(4):409–21.

434. Mehta A, Norton P, Liang H, Comunale MA, Wang M, Rodemich-Betesh L, et al. Increased levels of tetra-antennary N-linked glycan but not core fucosylation are associated with hepatocellular carcinoma tissue. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2012 Jun;21(6):925–33.

435. Block TM, Mehta AS, Fimmel CJ, Jordan R. Molecular viral oncology of hepatocellular carcinoma. Oncogene. 2003 Aug 11;22(33):5093–107.

436. El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med. 1999 Mar 11;340(10):745–50.

437. Okuda K, Fujimoto I, Hanai A, Urano Y. Changing incidence of hepatocellular carcinoma in Japan. Cancer Res. 1987 Sep 15;47(18):4967–72.

438. Taylor-Robinson SD, Foster GR, Arora S, Hargreaves S, Thomas HC. Increase in primary liver cancer in the UK, 1979-94. Lancet. 1997 Oct 18;350(9085):1142–3.

439. Deuffic S, Poynard T, Buffat L, Valleron AJ. Trends in primary liver cancer. Lancet. 1998 Jan 17;351(9097):214–5.

440. Chignard N, Beretta L. Proteomics for hepatocellular carcinoma marker discovery. Gastroenterology. 2004 Nov;127(5 Suppl 1):S120–5.

441. Zeindl-Eberhart E, Haraida S, Liebmann S, Jungblut PR, Lamer S, Mayer D, et al. Detection and identification of tumor-associated protein variants in human hepatocellular carcinomas. Hepatol Baltim Md. 2004 Feb;39(2):540–9.

442. Yokoo H, Kondo T, Fujii K, Yamada T, Todo S, Hirohashi S. Proteomic signature corresponding to alpha fetoprotein expression in liver cancer cells. Hepatol Baltim Md. 2004 Sep;40(3):609–17.

443. Melle C, Kaufmann R, Hommann M, Bleul A, Driesch D, Ernst G, et al. Proteomic profiling in microdissected hepatocellular carcinoma tissue using ProteinChip technology. Int J Oncol. 2004 Apr;24(4):885–91.

444. Qiao B, Wang J, Xie J, Niu Y, Ye S, Wan Q, et al. Detection and identification of peroxiredoxin 3 as a biomarker in hepatocellular carcinoma by a proteomic approach. Int J Mol Med. 2012 May;29(5):832–40.

445. Ai J, Tan Y, Ying W, Hong Y, Liu S, Wu M, et al. Proteome analysis of hepatocellular carcinoma by laser capture microdissection. Proteomics. 2006 Jan;6(2):538–46.

446. Kim J, Ki SS, Lee SD, Han CJ, Kim YC, Park SH, et al. Elevated plasma osteopontin levels in patients with hepatocellular carcinoma. Am J Gastroenterol. 2006 Sep;101(9):2051–9.

447. Chen R-X, Xia Y-H, Cui J-F, Xue T-C, Ye S-L. Osteopontin, a single marker for predicting the prognosis of patients with tumor-node-metastasis stage I hepatocellular carcinoma after surgical resection. J Gastroenterol Hepatol. 2010 Aug;25(8):1435–42.

448. Shang S, Plymoth A, Ge S, Feng Z, Rosen HR, Sangrajrang S, et al. Identification of osteopontin as a novel marker for early hepatocellular carcinoma. Hepatol Baltim Md. 2012 Feb;55(2):483–90.

449. Buamah PK, Cornell C, Cassells-Smith AJ, Harris AL. Fucosylation of alpha-fetoprotein in hepatocellular carcinomas. Lancet. 1986 Apr 19;1(8486):922–3.

450. Aoyagi Y. Molecular discrimination between alpha-fetoprotein from patients with hepatocellular-carcinoma and nonneoplastic liver-diseases by their carbohydrate structures (review). Int J Oncol. 1994 Feb;4(2):369–83.

451. Naitoh A, Aoyagi Y, Asakura H. Highly enhanced fucosylation of serum glycoproteins in patients with hepatocellular carcinoma. J Gastroenterol Hepatol. 1999 May;14(5):436–45.

452. Comunale MA, Lowman M, Long RE, Krakover J, Philip R, Seeholzer S, et al. Proteomic analysis of serum associated fucosylated glycoproteins in the development of primary hepatocellular carcinoma. J Proteome Res. 2006 Feb;5(2):308–15.

453. Comunale MA, Rodemich-Betesh L, Hafner J, Wang M, Norton P, Di Bisceglie AM, et al. Linkage Specific Fucosylation of Alpha-1-Antitrypsin in Liver Cirrhosis and Cancer Patients: Implications for a Biomarker of Hepatocellular Carcinoma. Ryu W-S, editor. PLoS ONE. 2010 Aug 25;5(8):e12419.

454. Comunale MA, Wang M, Hafner J, Krakover J, Rodemich L, Kopenhaver B, et al. Identification and development of fucosylated glycoproteins as biomarkers of primary hepatocellular carcinoma. J Proteome Res. 2009 Feb;8(2):595–602.

455. Marrero JA, Romano PR, Nikolaeva O, Steel L, Mehta A, Fimmel CJ, et al. GP73, a resident Golgi glycoprotein, is a novel serum marker for hepatocellular carcinoma. J Hepatol. 2005 Dec;43(6):1007–12.

456. Liu X-E, Desmyter L, Gao C-F, Laroy W, Dewaele S, Vanhooren V, et al. N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus. Hepatol Baltim Md. 2007 Nov;46(5):1426–35.

457. Goldman R, Ressom HW, Varghese RS, Goldman L, Bascug G, Loffredo CA, et al. Detection of hepatocellular carcinoma using glycomic analysis. Clin Cancer Res Off J Am Assoc Cancer Res. 2009 Mar 1;15(5):1808–13.

458. An HJ, Kronewitter SR, de Leoz MLA, Lebrilla CB. Glycomics and disease markers. Curr Opin Chem Biol. 2009 Dec;13(5-6):601–7.

459. Powers TW, Neely BA, Shao Y, Tang H, Troyer DA, Mehta AS, 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):e106255.

460. Hamm G, Bonnel D, Legouffe R, Pamelard F, Delbos J-M, Bouzom F, et al. Quantitative mass spectrometry imaging of propranolol and olanzapine using tissue extinction calculation as normalization factor. J Proteomics. 2012 Aug 30;75(16):4952–61.

461. Pirman DA, Reich RF, Kiss A, Heeren RMA, Yost RA. Quantitative MALDI Tandem Mass Spectrometric Imaging of Cocaine from Brain Tissue with a Deuterated Internal Standard. Anal Chem. 2013 Jan 15;85(2):1081–9.

462. Matos LL de, Trufelli DC, de Matos MGL, da Silva Pinhal MA. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. Biomark Insights. 2010;5:9–20.

463. Taniuchi K, Cerny RL, Tanouchi A, Kohno K, Kotani N, Honke K, et al. Overexpression of GalNAc-transferase GalNAc-T3 promotes pancreatic cancer cell growth. Oncogene. 2011 Dec 8;30(49):4843–54.

464. Mas E, Pasqualini E, Caillol N, El Battari A, Crotte C, Lombardo D, et al. Fucosyltransferase activities in human pancreatic tissue: comparative study between cancer tissues and established tumoral cell lines. Glycobiology. 1998 Jun;8(6):605–13.

465. Bassagañas S, Allende H, Cobler L, Ortiz MR, Llop E, de Bolós C, et al. Inflammatory cytokines regulate the expression of glycosyltransferases involved in the biosynthesis of tumor-associated sialylated glycans in pancreatic cancer cell lines. Cytokine. 2015 Apr 28;

466. Campbell MP, Royle L, Radcliffe CM, Dwek RA, Rudd PM. GlycoBase and autoGU: tools for HPLC-based glycan analysis. Bioinformatics. 2008 May 1;24(9):1214–6.

467. Biskup K, Braicu EI, Sehouli J, Fotopoulou C, Tauber R, Berger M, et al. Serum glycome profiling: a biomarker for diagnosis of ovarian cancer. J Proteome Res. 2013 Sep 6;12(9):4056–63.

468. Gustafsson OJR, Briggs MT, Condina MR, Winderbaum LJ, Pelzing M, McColl SR, et al. MALDI imaging mass spectrometry of N-linked glycans on formalin-fixed paraffin-embedded murine kidney. Anal Bioanal Chem. 2014 Dec 2;

469. Toghi Eshghi S, Yang S, Wang X, Shah P, Li X, Zhang H. Imaging of N-Linked Glycans from Formalin-Fixed Paraffin-Embedded Tissue Sections Using MALDI Mass Spectrometry. ACS Chem Biol. 2014 Sep 19;9(9):2149–56.

470. Pinsky, Claire S. Zhu. Building Multi-Marker Algorithms for Disease Prediction—The Role of Correlations Among Markers. Biomark Insights. 2011 Aug;83.

471. Zhang H, Chan DW. Cancer Biomarker Discovery in Plasma Using a Tissue-targeted Proteomic Approach. Cancer Epidemiol Biomark Amp Prev. 2007 Oct 1;16(10):1915–7.

472. Tep S, Hincapie M, Hancock WS. A general approach for the purification and quantitative glycomic analysis of human plasma. Anal Bioanal Chem. 2012 Mar;402(9):2687–700.

473. Stumpo KA, Reinhold VN. The N-Glycome of Human Plasma. J Proteome Res. 2010 Sep 3;9(9):4823–30.

474. Radhakrishnan P, Dabelsteen S, Madsen FB, Francavilla C, Kopp KL, Steentoft C, et al. Immature truncated O-glycophenotype of cancer directly induces oncogenic features. Proc Natl Acad Sci U S A. 2014 Sep 30;111(39):E4066–75.

475. Balog CIA, Stavenhagen K, Fung WLJ, Koeleman CA, McDonnell LA, Verhoeven A, et al. N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. Mol Cell Proteomics MCP. 2012 Sep;11(9):571–85.

476. Yue T, Partyka K, Maupin KA, Hurley M, Andrews P, Kaul K, et al. Identification of blood-protein carriers of the CA 19-9 antigen and characterization of prevalence in pancreatic diseases. Proteomics. 2011 Sep;11(18):3665–74.

477. Chandler K, Goldman R. Glycoprotein disease markers and single protein-omics. Mol Cell Proteomics MCP. 2013 Apr;12(4):836–45.

478. Yue T, Maupin KA, Fallon B, Li L, Partyka K, Anderson MA, et al. Enhanced discrimination of malignant from benign pancreatic disease by measuring the CA 19-9 antigen on specific protein carriers. PloS One. 2011;6(12):e29180.

479. Devakumar A, Mechref Y, Kang P, Novotny M, Reilly J. Identification of Isomeric N-Glycan Structures by Mass Spectrometry with 157 nm Laser-Induced Photofragmentation. J Am Soc Mass Spectrom. 2008 Jul;19(7):1027–40.

480. Domann P, Spencer DIR, Harvey DJ. Production and fragmentation of negative ions from neutral N-linked carbohydrates ionized by matrix-assisted laser desorption/ionization: Negative ions from neutral N-linked carbohydrates ionized by MALDI. Rapid Commun Mass Spectrom. 2012 Feb 29;26(4):469–79.

481. Smith DF, Kharchenko A, Konijnenburg M, Klinkert I, Paša-Tolić L, Heeren RMA. Advanced mass calibration and visualization for FT-ICR mass spectrometry imaging. J Am Soc Mass Spectrom. 2012 Nov;23(11):1865–72.

482. Tao S, Orlando R. A Novel Method for Relative Quantitation of N-Glycans by Isotopic Labeling Using 18O-Water. J Biomol Tech JBT. 2014 Dec;jbt.14–2504 – 003.

483. Walker SH, Taylor AD, Muddiman DC. The use of a xylosylated plant glycoprotein as an internal standard accounting for *N* -linked glycan cleavage and sample preparation variability: Plant glycoprotein internal standard. Rapid Commun Mass Spectrom. 2013 Jun 30;27(12):1354–8.

484. Fowler CB, O'Leary TJ, Mason JT. Toward improving the proteomic analysis of formalin-fixed, paraffin-embedded tissue. Expert Rev Proteomics. 2013 Aug;10(4):389–400.

485. Kobayashi Y, Tateno H, Dohra H, Moriwaki K, Miyoshi E, Hirabayashi J, et al. A Novel Core Fucose-specific Lectin from the Mushroom Pholiota squarrosa. J Biol Chem. 2012 Oct 5;287(41):33973–82.

486. Herreros-Villanueva M. Mouse models of pancreatic cancer. World J Gastroenterol. 2012;18(12):1286.

487. Schreiber SC, Giehl K, Kastilan C, Hasel C, Mühlenhoff M, Adler G, et al. Polysialylated NCAM represses E-cadherin-mediated cell-cell adhesion in pancreatic tumor cells. Gastroenterology. 2008 May;134(5):1555–66.

488. Fukasawa M, Matsushita A, Korc M. Neuropilin-1 interacts with integrin beta1 and modulates pancreatic cancer cell growth, survival and invasion. Cancer Biol Ther. 2007 Aug;6(8):1173–80.

489. Chang Z-G, Wei J-M, Qin C-F, Hao K, Tian X-D, Xie K, et al. Suppression of the Epidermal Growth Factor Receptor Inhibits Epithelial–Mesenchymal Transition in Human Pancreatic Cancer PANC-1 Cells. Dig Dis Sci. 2012 May;57(5):1181–9.

490. Ouyang H, Mou L, Luk C, Liu N, Karaskova J, Squire J, et al. Immortal Human Pancreatic Duct Epithelial Cell Lines with Near Normal Genotype and Phenotype. Am J Pathol. 2000 Nov;157(5):1623–31.

BIOGRAPHY

Thomas Wesley Powers, the son of Drs. David and Nancy Powers, was born in Ashville, NC in August of 1989. He has one older sister, Dr. Carolyn Elizabeth Powers, who was born 18 months prior. In his youth, Thomas was heavily involved in extracurricular and leadership activities, and was constantly encouraged to succeed academically. Following graduation from St. Joseph's Catholic School in 2007, Thomas pursued a Bachelor's of Science in chemistry at Furman University. During this time, he participated in undergraduate research in the analytical chemistry lab run by Dr. Wheeler, which provided him his first exposure to both mass spectrometry and research. Following his second year at Furman, Thomas developed a passion for economics and decided to obtain an additional Bachelor's of Arts in economics. To best combine these two divergent interests, Thomas decided to focus on applied research related to the health care industry at the Medical University of South Carolina. He enrolled and joined a research lab immediately following graduation from Furman University in May of 2011. Following lab rotations, Thomas ultimately decided to join Dr. Richard Drake's laboratory, where he continued to build upon his fundamentals in mass spectrometry, while focusing on addressing biological problems. During this time, Thomas published first author publications directed at localizing N-linked glycans using mass spectrometry imaging methods, and was asked to present at national conferences on the methodology. During his graduate school education, Thomas married Virginia Haywood Pazdan (June 30th, 2012) and had a daughter, Virginia Avery Powers (March 18th, 2015). Thomas successfully defended his Ph.D. dissertation in May of 2015.