A UNIQUE GLYCAN IS A SPECIFIC MARKER FOR PANCREATIC ADENOCARCINOMA

by

SAMUEL JOSEPH DOLEZAL

(Under the Direction of J. Michael Pierce)

ABSTRACT

Pancreatic adenocarcinoma (PDAC) is one of the most deadly cancers with one-year and five-year survival rates of only 24% and 5%, respectively. Physicians currently lack useful biomarkers in the screening, diagnosis, and treatment of PDAC, which unfortunately leads to the majority of patients being diagnosed in incurable, progressive stages or after metastasis has already occurred. Glycan structures expressed on glycoproteins are dynamic; quantitative and qualitative alternation of glycan expression is a hallmark of oncogenesis. These glycosylation changes are found on cell surface glycoconjugates and, in one case, has been specifically targeted as a therapeutic strategy. To identify a cancer-specific glycan target for PDAC, a mouse monoclonal IgG antibody (MAb109) has been identified that reacts with PDAC, but not with non-diseased pancreatic tissues. The epitope of this antibody is an N-glycan-specific epitope expressed on a small set of glycoproteins, as demonstrated by N-Glycosidase F (PNGaseF) sensitivity on immunoblots following SDS-PAGE. The specific glycoprotein in PDAC that expresses the epitope is CEACAM6, CD66c. Mass spectrometric analysis (MS) analysis performed on MAb-reactive CEACAM6 tryptic glycopeptides identified the epitope as an additional HexNAc attached to the β-linked core mannose, which was expressed only on one out of 12 N-linked sites. Additionally, truncation and site-directed mutagenesis identified three consecutive amino acids toward the C-terminus of CEACAM6 that function as a sequence determinant required for addition of the glyan epitope at its upstream glycosylation site. This determinant explains the stringent specificity of the epitope for expression on CEACAM6 and, in some cancers, CEACAM5, which has significant sequence identity with CEACAM6. MAb109 that targets this unique glycan epitope may demonstrate potential as a therapeutic for PDAC, while the epitope, which is released from PDAC cells, may serve as a potential diagnostic for this disease.

A UNIQUE GLYCAN IS A SPECIFIC MARKER FOR PANCREATIC ADENOCARCINOMA

by

SAMUEL JOSEPH DOLEZAL

B.S., The University of Georgia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

A UNIQUE GLYCAN IS A SPECIFIC MARKER FOR PANCREATIC ADENOCARCINOMA

by

SAMUEL JOSEPH DOLEZAL

Major Professor: Michael Pierce

Committee: Steven Dalton
Kojo Mensa Wilmot
Michael Tiemeyer

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
May 2015

DEDICATION

I dedicate this dissertation to my caring parents Timothy and Mavis who worked diligently, so that I would always have the best opportunity to succeed in my personal and professional endeavors.

Thanks, also to my loving wife, Shannon. Without you, I do not think that I would have ever made it.

ACKNOWLEDGEMENTS

It is with immense gratitude that I acknowledge the support and help of my Professor J. Michael Pierce for his constant guidance and support over the last five years. I cannot imagine a better advisor to guide and challenge me professionally, but also care about me personally. Thanks.

I am also indebted to my many colleagues in the Complex Carbohydrate Research Center who supported me in every way imaginable. This is a truly inspiring place to train in the field of glycobiology. I would like to express my gratitude to Kazuhiro Aoki, John Glushka, Kelley Moremen, Heather Strachan Moniz, Heather Steet, and Stephanie Stalnaker, and the many others I have not named whom all have helped me immensely with their expertise.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW AND INTRODUCTION

INTRODUCTION TO GLYCOSYLATION

Glycosylation is a post-translational modification involving the complex addition of oligosaccharides or polysaccharides to various aglycones, including to N-linked or O-linked sites on proteins. The glycans found in humans are predominantly attached to a protein either through an asparagine (N) residue contained within a N-X-S/T consensus sequence, where X is any amino acid except proline, or through the hydroxyl group of a serine or threonine residue [1]. The glycans attached to proteins exhibit a vast amount of structural diversity. This diversity is created primarily from the complexity of the monosaccharides, their sequence and possible linkages [2]. In addition to glycan diversity, protein molecules with the same polypeptide sequence can have distinct glycans attached at different sites [3]. Although the numbers and types of glycan structures attached to any protein are theoretically vast, constraints are provided by the mechanisms of glycan biosynthesis and its regulation.

Glycan biosynthesis predominantly occurs within the secretory pathway of the cell through the activities of glycosyltransferases and glycosidases. Glycosyltransferases are enzymes responsible for catalyzing the addition of a sugar from a high-energy sugar nucleotide to a specific substrate, while glycosidases catalyze the hydrolysis of glycosidic bonds to remove sugars [4, 5]. The majority of glycosyltransferases and glycosidases act sequentially, so the product of one enzyme is the preferred acceptor

substrate for the next enzyme. Additionally, glycosyltransferases use a specific donor substrate to catalyze a specific linkage [6, 7]. For example, N-acetylglucosaminyltransferase (GlcNAcT)-II uses UDP-N-acetylglucosamine (GlcNAc) to catalyze the β1,2 addition of GlcNAc to the α1,6 mannose of an N-linked glycan, but this specific addition only occurs when the previous glycosidase, α mannosidase II, has acted to reveal the preferred substrate product for GlcNAcT-II [8]. The underlying glycan structure, not the underlying polypeptide or glycolipid substrate, determines the specificity of the majority of glycosyltransferases and glycosidases. Although rare, examples of glycosylatransferases recognizing protein sequence determinants have been observed, including the polysialylation of the neural cell adhesion molecule (NCAM) [9, 10], O-fucosylation of thrombospondin repeats (TSRs) [11, 12], and GlcNAc-1-phosphotransferase modification of lysosomal enzymes [13].

Glycans are transferred to proteins and other sugars in the endoplasmic reticulum (ER) and Golgi, after which the glycoproteins are then commonly transported to the cell surface, endosomes, or lysosomes where they perform numerous cellular functions. These functions include linking proteins to cell membranes through GPI-anchors [14], mediating cell-cell interactions [15], functioning as signaling molecules [16], and determining the folding/stability/half-life of proteins [16]. Alterations in glycan structures can affect glycoprotein function, and altered glycans and glycoproteins are observed in cancer cells and tissues.

CANCER GLYCOMICS

Although glycan biosynthesis is regulated, evidence has shown that glycan structures are dynamic and commonly altered in diseases processes, due to the dysregulation of glycosylation enzyme transcription [17]. Because gene expression and cell signaling are altered in cancer cells, glycan structures are also affected. As a result, altered glycosylation is a hallmark of oncogenic transformation. The over-expression of the glycosyltransferase GnT-V in tumor cells is an example of altered cell signaling affecting cancer-specific glycosylation. Activation of the ras-raf-ets signaling pathway by her-2 increases GnT-V expression [18, 19]. This overexpression results in increased β1-6 branching of N-linked glycans that has been associated with increased L-PHA binding, extension of structures by poly-N-actyllactosamine, and increased invasiveness [20]. Additionally, the majority of human carcinomas exhibit the Tn-antigen, a GalNAc residue linked to a serine/threonine residue, which is normally extended by galactosyltransferase to produce the Core 2 T-antigen [21]. In carcinomas, a mutation in the molecular chaperone COSMC causes the galactosyl-transferase T-synthase to be misfolded, thus inactive, leading to an abundance of Tn-antigen [22]. Cancer cells appear to select for the COSMC mutation since Tn-antigen on the TRAIL-2 receptor causes aberrant signaling, allowing the cells to evade apoptosis [23]. Additional glycosylation changes associated with cancer also include over-expression of Sialyl-Lewis^x/Sialyl-Lewis^y structures [24], loss of GPI anchors [25], and increased incorporation of Neu5Gc into sialic acid glycoconjugates [26].

Because of the prevalence of glycosylation changes in cancer, the investigation of glycans and glycoproteins has now become increasingly important in the early diagnosis of cancer [27, 28]. Currently, nine out of ten of the most commonly used antibodies in clinical oncology as disease markers bind to glycans or glycoproteins [29]. Glycans and glycoproteins are useful markers because changes in glycosylation often occur during the early stages of oncogenesis, and they may also serve as a cell surface molecule for diagnosing and treating cancers [30]. Treatments targeting glycans could also prove effective since cells expressing these altered glycans commonly undergo increased cell proliferation and decrease cellular adhesions favoring metastasis; therefore, targeting these changes with a cytotoxic monoclonal antibody could potentially lead to the reduction of tumor growth and metastasis.

This dissertation focuses on the characterization of a monoclonal antibody (MAb), MAb109, and its binding epitope expressed on glycoproteins in cancer tissues, specifically in pancreatic cancer. Pancreatic cancer often goes undiagnosed until late stages and does not have an effective treatment largely as a result of the lack of an effective biomarker for the disease. Identification of a cancer-specific epitope will likely provide a biomarker specific to pancreatic cancer that would allow for the early detection of the disease and potentially lead to the development of effective therapeutics improving patient outcomes.

PANCREATIC CANCER

According to The American Cancer Society, 43,920 people were diagnosed with pancreatic cancer in the United States in 2012 [31]. Despite being the tenth most

common malignancy, pancreatic cancer will cause nearly forty thousand deaths annually making it the forth most deadly malignancy trailing behind only lung, colon, and breast cancers. Currently, the one-year and five-year survival rates of pancreatic cancer are 24% and 5%, respectively, amongst the worse for any malignancy [32]. While many of the most common cancers have recently experience increased survival rates, pancreatic cancer has exhibited the same rates for the past thirty years largely due to a lack of better screening and early detection methods [33]. Chances of developing better detection markers for pancreatic cancer increases as pancreatic function and precursor lesions become increasingly well characterized.

PANCREAS FUNCTION

The pancreas is a lobulated, retroperitoneal organ having been developed from the ventral and dorsal buds of the embryonic foregut, an endoderm derivative [34]. Stages of subsequent pancreatic development lead to the eventual formation of two distinct functional cell populations, endocrine and exocrine cells. The endocrine pancreas comprises the minority of the mass of the pancreas, only 1-2%, and is carefully organized into cellular structures known as the Islets of Langerhans [35]. These islets contain α, β, δ, and PP cells that secret glucagon, insulin, somatostatin, and pancreatic polypeptide, respectively, which are hormones predominantly involved in blood glucose homeostasis. Despite the importance of these cells to glucose homeostatis, the Islet cell types rarely develop cancer. In fact, approximately only 2% of all pancreatic neoplasms are islet cell tumors with the majority being benign. The majority of the pancreas, nearly 99% of the mass, is composed of acinar cells, larger ducts, and

smaller ductules, which together are responsible for the production and transport of digestive enzymes from the pancreas to the duodenum [36]. Because the epithelial cells lining the ducts and ductules of the pancreas undergo cell divisions to maintain exocrine function, this cell population is most commonly responsible for the acquisition and accumulation of mutations leading to the eventual development of pancreatic ductal adenocarcinoma.

PANCREATIC DUCTAL ADENOCARCINOMA (PDAC) PRECURSOR LESIONS

Previous studies have established that PDAC develops through a series of well-characterized, preinvasive stages during which curative treatment is possible if detected very early [37, 38]. The cell morphology, histology, and genetic mutations of these precursor lesions predictively progress to closely mimic that of the invasive carcinoma, and this progression from non-neoplastic lesion to carcinoma is more well-characterized in pancreatic cancer than in other cancers [39]. Currently, three PDAC precursor lesions have been identified-- pancreatic intraepithelial neoplasia (PanIN), intraductal papillary mucinous neoplasms (IPMNs), and mucinous cystic neoplasia (MCNs) [37, 40-42]. The predominant PDAC precursor lesion is pancreatic intraepithelial neoplasia (PanIN), accounting for over 90% of all eventual cases of PDAC.

Pancreatic intraepithelial neoplasia (PanINs)

Pancreatic intraepithelial neoplais (PanIN) represents the most common and well-characterized precursor lesion of PDAC. These lesions are defined as microscopic duct lesions that display similar cytohistological features as PDAC. PanINs represent progressive neoplastic growth as columnar epithelial cells replace the normal, flat

cuboidal epithelial lining of the pancreatic ducts [37]. The lesions are classified into PanIN-1A or PanIN-1B, PanIN-2, and PanIN-3 subtypes depending on the degree of cytological and architectural atypia. A low-grade PanIN, like PanIN-1, demonstrates little atypia or dysplasia whereas a high-grade PanIN, like PanIN-3, exhibits neoplasia with pervasive cellular atypia and dysplasia.

Despite the scientific evidence showing progression of pancreatic cancer occurs through histologically and molecularly well-characterized precursor lesions, the diagnosis and treatment of the disease has not improved dramatically over the years. The molecular transformation responsible for pancreatic cancer has become increasingly well characterized, but the clinical diagnoses of pancreatic cancer remains challenging for clinicians. The persistent delayed or missed diagnosis of pancreatic cancer can be attributed to the absence of early stage symptoms, the presence of vague, non-specific late stage symptoms, and the lack of efficient non-invasive diagnostic tests in clinical practice .

IMPORTANCE OF NEW PANCREATIC ADENOCARCINOMA BIOMARKERS

Pancreatic cancer biomarkers are needed because recent studies have shown that the cancer develops through a series of well-defined precursor lesions over time during which treatment is possible. The time pancreatic cancer takes to develop from initial mutations to metastasize was previously unknown; however, recent studies of the genetic evolution of pancreatic cancer tumors have shown that at least a decade passes between the initiating mutation in a normal duct epithelial cell and the

development of a clonal, non-metastatic parent tumor cell line [43-45]. On average, tumor cells require an additional five years to develop metastatic ability and another two after that before death occurs. This data defines a period of over fifteen years during which early detection and intervention could occur in order to improve clinical outcome of the disease and most importantly prevent patient deaths [43]. Studies predict that if pancreatic cancer could be detected, even in as little as six months before clinical presentation, then surgical resection could be performed and thus greatly impact patients' clinical outcomes.

Considering new data suggesting that a significant time window exists for detecting pancreatic cancers that arise from precursors while still in the curative stage, increased efforts into the discovery of biomarkers for early detection of this disease are essential in order to improve patient outcome. Improved screening and early detection methods depend largely on the development of sensitive and specific cancer markers, and the past decade has seen a plethora of advancements in the field of proteomics and glycomics, which coupled to the interest in early PDAC detection, is likely to lead to the identification of new, improved biomarkers replacing the existing ones.

BIOMARKERS CURRENTLY IN USE FOR PANCREATIC CANCER

A biomarker is defined as "a biological molecule found in blood, other bodily fluid, or tissues that is a sign of a normal or abnormal process or a condition of a disease" by the National Cancer Institute (NCI) [46]. Characteristics of a good cancer biomarker include the ability to diagnose cancer from non-diseased, normal tissue with a high sensitivity and specificity, are involved in the process that causes cancer or its progression, levels

should be high enough to measure easily, accurately, and reliably, and levels should correlate with the progression of the disease, and effective treatment should change the measured levels of the marker [47]. Additionally, blood-based biomarkers are preferred clinically because of their ability to be measured in an inexpensive, non-invasive blood draw.

Developing pancreatic cancer biomarkers could potentially help clinically to assess risk, aid in diagnosis, determine prognosis, predict treatment response, monitor treatment response, and detect recurrence of the disease [48]. Additionally, given the rarity of pancreatic cancer (incidence of 12.3 cases per 100,000), a test with a high sensitivity and specificity is required in order to keep a large population of individuals from unwarranted imaging or more invasive testing [49]. Better screening and early detection of pancreatic cancer depends largely on the development of improved cancer markers, but only one marker is used clinically with pancreatic cancer, CA19-9, a sialylated O-linked tetrasachharide [50].

CARBOHYDRATE ANTIGEN 19-9 (CA 19-9)

Serum carbohydrate antigen 19-9, also known as cancer antigen 19-9 or CA19-9, is a sialylated Lewis blood group antigen (SLe^A) with the structure, Siaα2,3Galβ1,3(Fucα1,4) GlcNAc [51]. CA19-9 was originally discovered as a tumor-specific antigen that reacted with a monoclonal antibody, N19.9, generated against colon cancer cells [52, 53]. Subsequent studies with the antibody demonstrated preferential reactivity with glycoproteins expressed the CA 19-9 glycan in pancreatic and biliary tract cancer patients, specifically mucins MUC1, MUC5, and MUC16 as well

as on ApoB, the primary apolipoprotein of chylomicrons and low-density lipoproteins [54, 55]. The CA19-9 antigen is currently the most commonly used marker clinically for the diagnosis of PDAC [56].

Because CA 19-9 is the standard serum tumor marker for PDAC, numerous studies have been published on its performance. In two large literature reviews summarizing all the studies, the authors found CA 19-9 to have a similar range of sensitivity and specificity, with an overall mean sensitivity of 80% and a mean specificity of 86% [50, 56]. The studies also determined that CA19-9 elevation in non-malignant jaundice resulted in a significant number of false positives, but these could be significantly reduced by using additional clinical tests. The main utility of CA 19-9 is, however, predicting clinical outcomes. Patients with elevated CA 19-9 levels that were unresponsive to adjuvant therapy had poorer survival rates compared to those normal or low levels [57-59].

Although CA 19-9 is the most widely used biomarker for pancreatic cancer, the marker has several drawbacks. For example, the specificity of CA 19-9 is affected by the fact that the epitope is observed in a variety of non-malignant diseases and additional cancer types. Benign diseases commonly found to have elevated CA 19-9 levels include biliary obstruction, cholecystitis, acute and chronic pancreatitis, cholangitis, obstructive jaundice, and liver cirrhosis. CA 19-9 levels are also increased in gastrointestinal tumors besides pancreatic adenocarcinoma including esophageal, gastric, colorectal, and bile duct cancers [60]. Additionally, CA 19-9 has poor sensitivity for the detection of small diameter pancreatic tumors with elevated levels observed in only 65% of patients with resectable pancreatic cancer [56]. Lastly, since CA 19-9 is a

sialyl lewis blood group antigen, the glycan is not produced, nor detected, in patients with the Lewis a-/b- phenotype. These patients lack the fucosyltransferase 3 required for the addition of the α1-4 fucose to GlcNAc of the CA19-9 tetrasaccharide [61]. As many as 10% of the Caucasian population have this phenotype [56]; therefore, this population tests negative for CA 19-9 even when tumor burden is high.

Because its drawbacks, CA 19-9 is currently only recommended as a measure of disease recurrence after surgical resection [62]. CA 19-9 levels decrease and remain low after successful surgical resection compared to measured preoperative levels; however, increases in postoperative levels suggest the recurrence of PDAC. When the CA 19-9 test is used diagnostically, clinical guidelines suggest that it should be used in conjunction with an imagining modality [63].

DEVELOPMENT OF MONOCLONAL ANTIBODY-BASED CANCER THERAPIES

Despite the continued use of CA 19-9 in the detection of pancreatic cancer, the morbidity and mortality of the disease remains high. Detection of pancreatic cancer continually occurs in the late stages when surgical resection is not possible. Additionally, the standard of care, Gemcitabine, a fluorinated nucleoside analog, with or without Erlotinib, a tyrosine kinase inhibitor of the epidermal growth factor receptor (EGFR), is commonly ineffective [64]. Pancreatic cancer lacks biomarker targets to take advantage of recent tumor-specific therapeutics, which are just now beginning to gain the specificity required to effectively treat patients.

Cancer therapeutics are experiencing a large movement towards the development of biological drugs based on monoclonal antibodies. World-wide sales of all therapeutic

monoclonal antibodies has risen to $57 billion in 2013 up from just $4 billion in 2001 [65]. Examples of monoclonal antibody-based therapies against cancer include trastuzumab, bevacizumab, and cetuximab, which specifically target the HER2/neu recepctor, vascular endothelial growth factor A (VEGF-A), and the epithelial growth factor receptor (EGFR), respectively [66-68]. The importance of monoclonal antibody based therapies in cancer is predicted to only further increase as more drugs enter clinical trials and are approved by the United States Food and Drug Administration (FDA) for clinical use [69]. Because of the success of these early antibody-based drugs, new classes of drugs, including antibody drug conjugates (ADCs), are being developed and appear poised to become the next-generation cancer therapeutic.

ADCs are monoclonal antibodies engineered to specifically deliver potent cytotoxic drugs to cancer cells. Two antibody-drug conjugates, ado-trastuzumab emtansine (Kadcyla®) and brentuximab vedotin (Avastin®), are currently approved by the FDA. Approved in 2011, brentuximab vedotin (Avastin®) is an anti-CD30 monoclonal antibody linked to an antimitotic agent, monomethyl auristatin E (MMAE), used to treat non-Hodgkin's lymphoma (NHL), and ado-trastuzumab emtansine (Kadcyla®) is a monoclonal antibody targeting HER2 delivering a cytoxic payload resulting in cancer cell death [70, 71]. Several additional ADCs are currently in clinical trials and will likely enter the market in the near future. These include CDX011 (glembatumumab vedotin), which targets glycoprotein non-metastatic melanoma protein B (GPNMB), pinatuzumab vedotin, which targets CD22 in lymphomas, and polatuzumab vedotin, which targets CD79b in NHL and diffuse large B cell lymphoma [72, 73]. ADCs are currently only

FDA-approved for blood tumors; however, ADCs directed against solid tumors including breast and prostate are in clinical trial Phase 2 or later [74].

The effectiveness of monoclonal antibodies in ADC-based cancer therapies depends largely on four factors, the antigen, antibody, linker, and cytotoxic payload, but the specificity towards the cancer cell depends solely on the antigen and antibody interaction [75]. Since ADCs are designed to bind and kill cancer cells in a target-dependent manner, tumor antigen recognition and binding by the antibody is crucial [76]. The tumor antigen must be expressed on the cell surface of tumor cells for recognition by the antibody and exhibit differential expression compared to non-diseased cells [69]. Differential antigen expression between tumor and non-diseased cells allows the specific targeting of tumor cells by the ADC while sparing non-diseased cells. Without differential expression of the antigen, non-diseased cells would bind and internalize the antibody, resulting in toxicity for non-diseased cells as well as reduced effective dose of ADC for tumor cells [77, 78]. Because of the low prevalence and difficulty identifying cancer-specific antigens, the rate-limiting step in ADC development is the identification of these antigens and the development of monoclonal antibodies towards them [79].

Cancer biomarkers based on glycans or glycoconjugates would provide exceptional targets for ADC. Altered glycosylation often occurs early during oncogenesis and generates cell-surface targets, which exhibit differential expression between non-diseased and cancer cells. Developing antibodies against these cancer-specific glycans would be useful in the diagnosis and treatment of cancer; however, generating

antibodies using the most common immunizing techniques does not always insure high-affinity antibodies which specifically bind glycan. Currently, monoclonal antibodies against glycans are generated by the immunization of mice with whole cells, isolated membrane preparations, or with a purified conjugate such as glycoprotein, glycolipid, or glycosaminoglycan (GAG). Antibody generation using these techniques results in the isolation of mostly IgM antibodies, which have higher valency but can affect binding specificity. IgG antibodies are preferred over IgM antibodies because of their specificity, but immunization with glycans or glycoconjugates does not always result in class switching. Effective generation of high-affinity monoclonal antibodies against specific cancer-specific glycans would not only likely identify a biomarker for the early detection of the disease, but also provide a platform from which new therapeutics could be developed.

DEVELOPMENT OF CANCER-SPECIFIC MONOCLONAL ANTIBODIES

A collaborator, Dr. Cohava Gelber, has generated a cancer-specific antibody that does not react with normal cells or tissues that have been tested [80]. The antibody, known as MAb109, was generated using an immunization technique known as Differential Immunization for Antigen and Antibody Discovery (DIADD). Briefly, mice were immunized with a neuroendocrine cell line to induced tolerance before being immunized a week later with a small cell lung cancer (SCLC) cell line. By inducing tolerance before immunological challenge with cancer cells, the DIAAD technique overcomes the bias of an immune response toward dominant, more prevalent antigens while generating an immune response against cancer antigens. MAb109 was isolated

through positive SCLC cell binding, but the antibody did not crossreact with closely related neuroendocrine cells.

Immunoblots of SCLC membrane antigens that were separated by SDS-PAGE under reducing conditions revealed MAb109 reactivity with a single, approximately 200 kDa band [80]. Interestingly, immunoblots of pancreatic cancer cells, another cell type tested for cross-reactivity, revealed strong antibody binding with a pattern distinct from SCLC. In addition to the 200 kDa band observed in SCLC, an additional, more prominent reactive band was observed at approximately 85 kDa. Furthermore, diminished antibody binding was observed following PNGaseF digestion. Taken together, these results suggested a common, shared epitope on distinct molecules expressed between SCLC and pancreatic cancer and that carbohydrates, specifically N-linked glycans, were involved in antibody recognition of the epitope.

Additionally, MAb109 immunohistochemical (IHC) staining of lung tissue sections showed increased antibody binding to epithelial cells from cancerous lung when compared to normal lung tissue and control IgG antibody. Because MAb109 demonstrated positive reactivity on immunoblots of pancreatic cancer cells and in IHC studies, next logical step was to perform IHC staining with the antibody on pancreatic tissue sections. A collaboration with Dr. Daniel Von Hoff and Haiyong Han of the Translational Genomics Research Institute (TGen) was established to perform the IHC staining, as well as calculate the sensitivity and specificity of MAb109. Sensitivity and specificity are common statistical measurements of the effectiveness of a biomarker at detecting the presence or absence of a disease. In our studies, the sensitivity would be the percentage of patients with pancreatic cancer that test positive while specificity

would be the percentage of non-diseased patients that are identified as not having pancreatic cancer.

MAB109 POSITIVELY STAINS PANCREATIC ADENOCARCINOMA TISSUE SECTIONS.

Our results showed MAb109 reactivity in pancreatic cancer tissue over non-diseased normal, and the staining was especially concentrated to the epithelial cells lining the ducts of pancreatic adenocarcinoma (Fig 1.1A). Additionally, less intense staining was also be observed in the spaces between the ducts, which typically undergo a desmoplastic response due to the infiltration of fibroblasts and inflammatory cells [81]. In contrast, the normal pancreas sections exhibited very little or no MAb109 staining, indicating that the normal tissue lacks the antigen recognized by the antibody. Even with higher magnification, the degree of MAb109 staining of normal tissue is negligible showing lack of expression of the antigen when compared to cancer cells (Fig. 1.1B). Additionally, Fig.1.1C was able to show that MAb109 stained dysplastic pancreas, which is a precursor to PDAC, as well as PDAC, and the degree of MAb109 corresponded to disease progression. From the greater than 150 tissue sections that have been analyzed thus far, the sensitivity and specificity has been calculated to be 98% and 82% respectively. These data showed that MAb109 reactive with early-stage pancreatic cancer, and the values were improvements over the current pancreatic cancer biomarker, CA19-9.

MAB109 REACTS WITH CARCINOEMBRYONIC ANTIGEN CELL ADHESION MOLECULE 6 (CEACAM6) IN PANCREATIC CANCER

The MAb109 reactive glycoepitope has since been identified by proteomic analysis and confirmed by immunoblotting to be expressed on carcinoembryonic antigen cell adhesion molecule 6 (CEACAM6) in BxPC3 cells, a pancreatic adenocarcinoma cell line. CEACAM6 is the predominant MAb109 reactive species likely as a result of its overexpression in pancreatic cancer, which is well characterized in the literature [82]. The CEACAM6 glycoprotein is a cell surface glycoprotein and member of the CEACAM family, which contains 12 genes and additional pseudogenes in humans [83]. CEACAM members are cell adhesion molecules characterized by a single amino-terminal immunoglobulin variable-like domain (IgV-like) followed by immunoglobulin constant-like domains, which may be attached to the cell surface utilizing either a transmembrane domain or glycophosphotidylinositol (GPI) anchor. CEACAM6 follows this pattern, expressing a single IgV-like domain, two IgC_2-like domains, and a GPI-anchor. Additionally, CEACAM6 contains twelve predicted N-linked glycosylation sites, which would account for a large mass difference between the predicted mass of the polypeptide and the observed mass of the glycoprotein upon SDS-PAGE [83]. CEACAM6 was originally discovered as an oncofetal antigen [84], an antigen generally expressed during development, downregulated during adolescence and adulthood, and expressed again during oncogenesis.

CEACAM6 is a tumor-associated antigen that is overexpressed in both primary and metastatic malignant pancreas [85, 86]. Staining of PDAC biopsy samples have shown CEACAM6 levels are approximately 20-25 higher when compared to non-diseased normal [85]. Additionally, the CEACAM6 expression level directly correlates with the

degree of tumor differentiation, and overexpression has been well-established role in cancer progression by affecting cell polarity, cell differentiation, and tissue architecture [87, 88]. Accumulating evidence also suggests that disruption in CEACAM6 expression is functionally important in cancer tumorigenesis, as demonstrated by changes in cell adhesion, cell invasiveness, resistance to anoikis, and metastatic behavior of tumor cells [88-91]. Increased CEACAM6 expression also correlates with decreased sensitivity to chemotherapeutic drugs like gemcitabine [92].

The goal of this dissertation research was to characterize the structure of the MAb109 epitope on CEACAM6 and investigate the specific addition of the epitope to a small subset of glycoproteins. The working hypothesis was that the MAb109 epitope is an unusual N-linked glycan structure that is expressed in pancreatic adenocarcinoma on proteins containing specific sequence determinants. By investigating the structure of this glycan and determinants required for its addition, we will very likely identify a unique cancer-specific glycan structure, provide a better understanding of protein-specific glycosylation, and set the stage to develop the antibody in the diagnosis and treatment of pancreatic cancer.

REFERENCES

1. Pless, D.D. and W.J. Lennarz, *Enzymatic conversion of proteins to glycoproteins.* Proc Natl Acad Sci U S A, 1977. **74**(1): p. 134-8.
2. Moremen, K.W., M. Tiemeyer, and A.V. Nairn, *Vertebrate protein glycosylation: diversity, synthesis and function.* Nat Rev Mol Cell Biol, 2012. **13**(7): p. 448-62.
3. Kamiya, Y., T. Satoh, and K. Kato, *Recent advances in glycoprotein production for structural biology: toward tailored design of glycoforms.* Curr Opin Struct Biol, 2014. **26C**: p. 44-53.
4. Bourne, Y. and B. Henrissat, *Glycoside hydrolases and glycosyltransferases: families and functional modules.* Curr Opin Struct Biol, 2001. **11**(5): p. 593-600.

5. Coutinho, P.M., et al., *An evolving hierarchical family classification for glycosyltransferases.* J Mol Biol, 2003. **328**(2): p. 307-17.
6. Ihara, H., Y. Ikeda, and N. Taniguchi, *Reaction mechanism and substrate specificity for nucleotide sugar of mammalian alpha1,6-fucosyltransferase--a large-scale preparation and characterization of recombinant human FUT8.* Glycobiology, 2006. **16**(4): p. 333-42.
7. Sarkar, M., et al., *Molecular cloning and expression of cDNA encoding the enzyme that controls conversion of high-mannose to hybrid and complex N-glycans: UDP-N-acetylglucosamine: alpha-3-D-mannoside beta-1,2-N-acetylglucosaminyltransferase I.* Proc Natl Acad Sci U S A, 1991. **88**(1): p. 234-8.
8. Bendiak, B. and H. Schachter, *Control of glycoprotein synthesis. Kinetic mechanism, substrate specificity, and inhibition characteristics of UDP-N-acetylglucosamine:alpha-D-mannoside beta 1-2 N-acetylglucosaminyltransferase II from rat liver.* J Biol Chem, 1987. **262**(12): p. 5784-90.
9. Thompson, M.G., D.A. Foley, and K.J. Colley, *The polysialyltransferases interact with sequences in two domains of the neural cell adhesion molecule to allow its polysialylation.* J Biol Chem, 2013. **288**(10): p. 7282-93.
10. Thompson, M.G., et al., *Sequences at the interface of the fifth immunoglobulin domain and first fibronectin type III repeat of the neural cell adhesion molecule are critical for its polysialylation.* J Biol Chem, 2011. **286**(6): p. 4525-34.
11. Harris, R.J. and M.W. Spellman, *O-linked fucose and other post-translational modifications unique to EGF modules.* Glycobiology, 1993. **3**(3): p. 219-24.
12. Panin, V.M., et al., *Notch ligands are substrates for protein O-fucosyltransferase-1 and Fringe.* J Biol Chem, 2002. **277**(33): p. 29945-52.
13. Steet, R., W.S. Lee, and S. Kornfeld, *Identification of the minimal lysosomal enzyme recognition domain in cathepsin D.* J Biol Chem, 2005. **280**(39): p. 33318-23.
14. Hwa, K.Y., *Glycosyl phosphatidylinositol-linked glycoconjugates: structure, biosynthesis and function.* Adv Exp Med Biol, 2001. **491**: p. 207-14.
15. Hildebrandt, H., et al., *Dissecting polysialic acid and NCAM functions in brain development.* J Neurochem, 2007. **103 Suppl 1**: p. 56-64.
16. Wells, L., S.A. Whelan, and G.W. Hart, *O-GlcNAc: a regulatory post-translational modification.* Biochem Biophys Res Commun, 2003. **302**(3): p. 435-41.
17. Ohtsubo, K. and J.D. Marth, *Glycosylation in cellular mechanisms of health and disease.* Cell, 2006. **126**(5): p. 855-67.
18. Buckhaults, P., et al., *Transcriptional regulation of N-acetylglucosaminyltransferase V by the src oncogene.* J Biol Chem, 1997. **272**(31): p. 19575-81.
19. Kang, R., et al., *Transcriptional regulation of the N-acetylglucosaminyltransferase V gene in human bile duct carcinoma cells (HuCC-T1) is mediated by Ets-1.* J Biol Chem, 1996. **271**(43): p. 26706-12.
20. Fernandes, B., et al., *Beta 1-6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia.* Cancer Res, 1991. **51**(2): p. 718-23.
21. Springer, G.F., *T and Tn, general carcinoma autoantigens.* Science, 1984. **224**(4654): p. 1198-206.

22. Ju, T., et al., *Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc.* Cancer Res, 2008. **68**(6): p. 1636-46.
23. Wagner, K.W., et al., *Death-receptor O-glycosylation controls tumor-cell sensitivity to the proapoptotic ligand Apo2L/TRAIL.* Nat Med, 2007. **13**(9): p. 1070-7.
24. Kannagi, R., *Molecular mechanism for cancer-associated induction of sialyl Lewis X and sialyl Lewis A expression-The Warburg effect revisited.* Glycoconj J, 2004. **20**(5): p. 353-64.
25. Chen, R., et al., *Glycophosphatidylinositol-anchored protein deficiency as a marker of mutator phenotypes in cancer.* Cancer Res, 2001. **61**(2): p. 654-8.
26. Malykh, Y.N., R. Schauer, and L. Shaw, *N-Glycolylneuraminic acid in human tumours.* Biochimie, 2001. **83**(7): p. 623-34.
27. Dube, D.H. and C.R. Bertozzi, *Glycans in cancer and inflammation--potential for therapeutics and diagnostics.* Nat Rev Drug Discov, 2005. **4**(6): p. 477-88.
28. Shriver, Z., S. Raguram, and R. Sasisekharan, *Glycomics: a pathway to a class of new and improved therapeutics.* Nat Rev Drug Discov, 2004. **3**(10): p. 863-73.
29. Pierce, J.M., *Cancer Glycomics. Handbook of Glycomics*, ed. Cummings. 2009, San Diego: Academic Press.
30. Kobata, A. and J. Amano, *Altered glycosylation of proteins produced by malignant cells, and application for the diagnosis and immunotherapy of tumours.* Immunol Cell Biol, 2005. **83**(4): p. 429-39.
31. *American Cancer Society*. Cancer Facts & Figures, 2014.
32. Jemal, A., et al., *Cancer statistics, 2009.* CA Cancer J Clin, 2009. **59**(4): p. 225-49.
33. Winter, J.M., et al., *Survival after resection of pancreatic adenocarcinoma: results from a single institution over three decades.* Ann Surg Oncol, 2012. **19**(1): p. 169-75.
34. Adda, G., L. Hannoun, and J. Loygue, *Development of the human pancreas: variations and pathology. A tentative classification.* Anat Clin, 1984. **5**(4): p. 275-83.
35. Cabrera, O., et al., *The unique cytoarchitecture of human pancreatic islets has implications for islet cell function.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2334-9.
36. Grapin-Botton, A., *Ductal cells of the pancreas.* Int J Biochem Cell Biol, 2005. **37**(3): p. 504-10.
37. Hruban, R.H., et al., *Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions.* Am J Surg Pathol, 2001. **25**(5): p. 579-86.
38. Hruban, R.H., et al., *Progression model for pancreatic cancer.* Clin Cancer Res, 2000. **6**(8): p. 2969-72.
39. Brat, D.J., et al., *Progression of pancreatic intraductal neoplasias to infiltrating adenocarcinoma of the pancreas.* Am J Surg Pathol, 1998. **22**(2): p. 163-9.
40. Salvia, R., et al., *Pancreatic cystic tumors.* Minerva Chir, 2004. **59**(2): p. 185-207.
41. Compagno, J. and J.E. Oertel, *Mucinous cystic neoplasms of the pancreas with overt and latent malignancy (cystadenocarcinoma and cystadenoma). A clinicopathologic study of 41 cases.* Am J Clin Pathol, 1978. **69**(6): p. 573-80.

42. Hruban, R.H., et al., *An illustrated consensus on the classification of pancreatic intraepithelial neoplasia and intraductal papillary mucinous neoplasms.* Am J Surg Pathol, 2004. **28**(8): p. 977-87.
43. Yachida, S., et al., *Distant metastasis occurs late during the genetic evolution of pancreatic cancer.* Nature, 2010. **467**(7319): p. 1114-7.
44. Campbell, P.J., et al., *The patterns and dynamics of genomic instability in metastatic pancreatic cancer.* Nature, 2010. **467**(7319): p. 1109-13.
45. Luebeck, E.G., *Cancer: Genomic evolution of metastasis.* Nature, 2010. **467**(7319): p. 1053-5.
46. Biomarkers Definitions Working, G., *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.* Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.
47. Etzioni, R., et al., *The case for early detection.* Nat Rev Cancer, 2003. **3**(4): p. 243-52.
48. Tanday, S., *Biomarkers in blood could help to detect pancreatic cancer.* Lancet Oncol, 2014. **15**(3): p. e108.
49. Surveillance, E., and End Results (SEER) Program, *SEER Research Data 1973-2011*, N.C. Institute, Editor. April 2014.
50. Goonetilleke, K.S. and A.K. Siriwardena, *Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer.* Eur J Surg Oncol, 2007. **33**(3): p. 266-70.
51. Hanisch, F.G., G. Uhlenbruck, and C. Dienst, *Structure of tumor-associated carbohydrate antigen Ca 19-9 on human seminal-plasma glycoproteins from healthy donors.* Eur J Biochem, 1984. **144**(3): p. 467-73.
52. Magnani, J.L., et al., *A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-N-fucopentaose II.* J Biol Chem, 1982. **257**(23): p. 14365-9.
53. Koprowski, H., et al., *Colorectal carcinoma antigens detected by hybridoma antibodies.* Somatic Cell Genet, 1979. **5**(6): p. 957-71.
54. Magnani, J.L., et al., *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. **43**(11): p. 5489-92.
55. Yue, T., et al., *Identification of blood-protein carriers of the CA 19-9 antigen and characterization of prevalence in pancreatic diseases.* Proteomics, 2011. **11**(18): p. 3665-74.
56. Ballehaninna, U.K. and R.S. Chamberlain, *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. **3**(2): p. 105-19.
57. Waraya, M., et al., *Preoperative serum CA19-9 and dissected peripancreatic tissue margin as determiners of long-term survival in pancreatic cancer.* Ann Surg Oncol, 2009. **16**(5): p. 1231-40.
58. Sperti, C., et al., *CA 19-9 as a prognostic index after resection for pancreatic cancer.* J Surg Oncol, 1993. **52**(3): p. 137-41.
59. Kondo, N., et al., *Prognostic impact of perioperative serum CA 19-9 levels in patients with resectable pancreatic cancer.* Ann Surg Oncol, 2010. **17**(9): p. 2321-9.

60. Kim, J.E., et al., *Clinical usefulness of carbohydrate antigen 19-9 as a screening test for pancreatic cancer in an asymptomatic population.* J Gastroenterol Hepatol, 2004. **19**(2): p. 182-6.
61. Kannagi, R., *Carbohydrate antigen sialyl Lewis a--its pathophysiological significance and induction mechanism in cancer progression.* Chang Gung Med J, 2007. **30**(3): p. 189-209.
62. Steinberg, W., *The clinical utility of the CA 19-9 tumor-associated antigen.* Am J Gastroenterol, 1990. **85**(4): p. 350-5.
63. Locker, G.Y., et al., *ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer.* J Clin Oncol, 2006. **24**(33): p. 5313-27.
64. Moore, M.J., et al., *Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group.* J Clin Oncol, 2007. **25**(15): p. 1960-6.
65. Visiongain, *Next-Generation Antibody Therapeutic Market 2013-2023.* 2013.
66. Hudis, C.A., *Trastuzumab--mechanism of action and use in clinical practice.* N Engl J Med, 2007. **357**(1): p. 39-51.
67. Los, M., J.M. Roodhart, and E.E. Voest, *Target practice: lessons from phase III trials with bevacizumab and vatalanib in the treatment of advanced colorectal cancer.* Oncologist, 2007. **12**(4): p. 443-50.
68. Messersmith, W.A. and D.J. Ahnen, *Targeting EGFR in colorectal cancer.* N Engl J Med, 2008. **359**(17): p. 1834-6.
69. Teicher, B.A., *Antibody-drug conjugate targets.* Curr Cancer Drug Targets, 2009. **9**(8): p. 982-1004.
70. Francisco, J.A., et al., *cAC10-vcMMAE, an anti-CD30-monomethyl auristatin E conjugate with potent and selective antitumor activity.* Blood, 2003. **102**(4): p. 1458-65.
71. Niculescu-Duvaz, I., *Trastuzumab emtansine, an antibody-drug conjugate for the treatment of HER2+ metastatic breast cancer.* Curr Opin Mol Ther, 2010. **12**(3): p. 350-60.
72. Keir, C.H. and L.T. Vahdat, *The use of an antibody drug conjugate, glembatumumab vedotin (CDX-011), for the treatment of breast cancer.* Expert Opin Biol Ther, 2012. **12**(2): p. 259-63.
73. Evans, J.B. and B.A. Syed, *Next-generation antibodies.* Nat Rev Drug Discov, 2014.
74. Lambert, J.M., *Drug-conjugated antibodies for the treatment of cancer.* Br J Clin Pharmacol, 2013. **76**(2): p. 248-62.
75. Flygare, J.A., T.H. Pillow, and P. Aristoff, *Antibody-drug conjugates for the treatment of cancer.* Chem Biol Drug Des, 2013. **81**(1): p. 113-21.
76. Panowksi, S., et al., *Site-specific antibody drug conjugates for cancer therapy.* MAbs, 2014. **6**(1): p. 34-45.
77. Senter, P.D., *Potent antibody drug conjugates for cancer therapy.* Curr Opin Chem Biol, 2009. **13**(3): p. 235-44.
78. Tolcher, A.W., et al., *Randomized phase II study of BR96-doxorubicin conjugate in patients with metastatic breast cancer.* J Clin Oncol, 1999. **17**(2): p. 478-84.

79. Beck, A., et al., *Strategies and challenges for the next generation of therapeutic antibodies.* Nat Rev Immunol, 2010. **10**(5): p. 345-52.
80. Krueger, P., et al., *A new small cell lung cancer (SCLC)-specific marker discovered through antigenic subtraction of neuroblastoma cells.* Cancer Immunol Immunother, 2003. **52**(6): p. 367-77.
81. Erkan, M., *The role of pancreatic stellate cells in pancreatic cancer.* Pancreatology, 2013. **13**(2): p. 106-9.
82. Duxbury, M.S., et al., *CEACAM6 is a novel biomarker in pancreatic adenocarcinoma and PanIN lesions.* Ann Surg, 2005. **241**(3): p. 491-6.
83. Hammarstrom, S., *The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues.* Semin Cancer Biol, 1999. **9**(2): p. 67-81.
84. von Kleist, S., G. Chavanel, and P. Burtin, *Identification of an antigen from normal human tissue that crossreacts with the carcinoembryonic antigen.* Proc Natl Acad Sci U S A, 1972. **69**(9): p. 2492-4.
85. Blumenthal, R.D., et al., *Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers.* BMC Cancer, 2007. **7**: p. 2.
86. Scholzel, S., et al., *Carcinoembryonic antigen family members CEACAM6 and CEACAM7 are differentially expressed in normal tissues and oppositely deregulated in hyperplastic colorectal polyps and early adenomas.* Am J Pathol, 2000. **156**(2): p. 595-605.
87. Poola, I., et al., *Expression of carcinoembryonic antigen cell adhesion molecule 6 oncoprotein in atypical ductal hyperplastic tissues is associated with the development of invasive breast cancer.* Clin Cancer Res, 2006. **12**(15): p. 4773-83.
88. Blumenthal, R.D., H.J. Hansen, and D.M. Goldenberg, *Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen).* Cancer Res, 2005. **65**(19): p. 8809-17.
89. Ordonez, C., et al., *Human carcinoembryonic antigen functions as a general inhibitor of anoikis.* Cancer Res, 2000. **60**(13): p. 3419-24.
90. Duxbury, M.S., et al., *CEACAM6 is a determinant of pancreatic adenocarcinoma cellular invasiveness.* Br J Cancer, 2004. **91**(7): p. 1384-90.
91. Duxbury, M.S., et al., *CEACAM6 gene silencing impairs anoikis resistance and in vivo metastatic ability of pancreatic adenocarcinoma cells.* Oncogene, 2004. **23**(2): p. 465-73.
92. Duxbury, M.S., et al., *A novel role for carcinoembryonic antigen-related cell adhesion molecule 6 as a determinant of gemcitabine chemoresistance in pancreatic adenocarcinoma cells.* Cancer Res, 2004. **64**(11): p. 3987-93.
93. Pan, S., et al., *Mass spectrometry based glycoproteomics--from a proteomics perspective.* Mol Cell Proteomics, 2011. **10**(1): p. R110 003251.

Marker	Disease	Glycan/Glycoprotein
α-Fetoprotein	Nonseminomatous testicular cancer	Glycoprotein
Human chorionic gonadotropin (hCG)	Testicular cancer	Glycoprotein
Carcinoembroyinc antigen (CEA)	Colon cancer	Glycoprotein
Prostate specific antigen (PSA)	Prostate cancer	Glycoprotein
Thyroglobulin	Thyroid cancer	Glycoprotein
CA 19-9	Pancreatic cancer	Glycan
CA125	Ovarian cancer	Glycan
CA 15-3	Breast cancer	Glycan
CA 27-29	Breast cancer	Glycan

Table 1.1. UNITED STATES FOOD AND DRUG APPROVED CANCER BIOMARKERS DETECTING GLYCANS OR GLYCOPROTEINS.

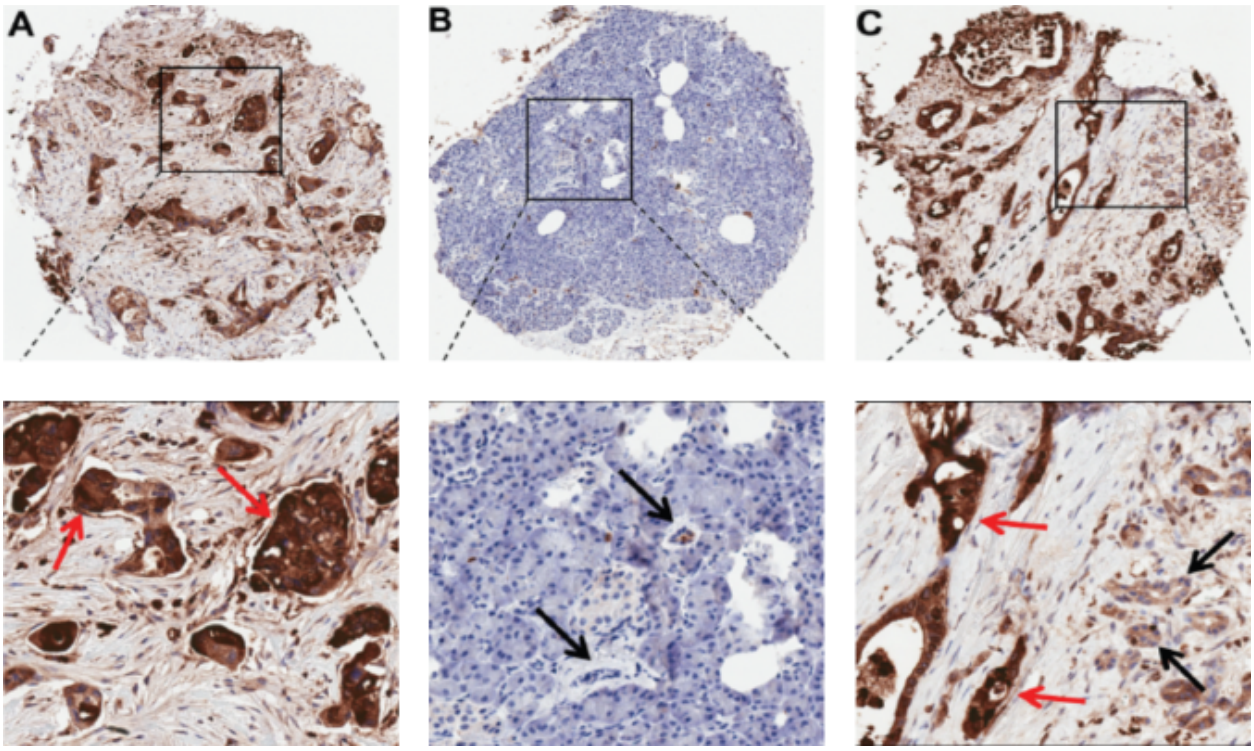


Figure 1.1. IMMUNOHISTOCHEMICAL STAINING OF PDAC SECTIONS WITH MAB109. (A) The antibody heavily stains the epithelial cells lining the pancreatic ducts (red arrows column A) in comparison to normal pancreas shown in column B (black arrows). Column C shows staining of both dysplastic pancreatic epithelium (black arrows) and PDAC (red arrows) with the staining intensity directly corresponding to the degree of disease progression.

CHAPTER 2: A UNIQUE GLYCAN IS A SPECIFIC MARKER FOR PANCREATIC ADENOCARCINOMA

Samuel J. Dolezal, Kazuhiro Aoki, John Glushka, Michael Tiemeyer, Cohava Gelber, and J. Michael Pierce. Manuscript to be submitted to Cancer Cell.

ABSTRACT

Pancreatic adenocarcinoma (PDAC) is one of the most deadly cancers with one-year and five-year survival rates of only 24% and 5%, respectively. Physicians currently lack useful biomarkers in the screening, diagnosis, and treatment of PDAC, which unfortunately leads to the majority of patients being diagnosed in incurable, progressive stages or after metastasis has already occurred. Glycan structures expressed on glycoproteins are dynamic; quantitative and qualitative alternation of glycan expression is a hallmark of oncogenesis. These glycosylation changes are found on cell surface glycoconjugates and, in one case, has been specifically targeted as a therapeutic strategy. To identify a cancer-specific glycan target for PDAC, a mouse monoclonal IgG antibody (MAb109) has been identified that reacts with PDAC, but not with non-diseased pancreatic tissues. The epitope of this antibody is an N-glycan-specific epitope expressed on a small set of glycoproteins, as demonstrated by N-Glycosidase F (PNGaseF) sensitivity on immunoblots following SDS-PAGE. The specific glycoprotein in PDAC that expresses the epitope is CEACAM6, CD66c. Mass spectrometric analysis (MS) analysis performed on MAb-reactive CEACAM6 tryptic glycopeptides identified the epitope as an additional HexNAc attached to the β-linked core mannose, which was expressed only on one out of 12 N-linked sites. Additionally, truncation and site-directed mutagenesis identified three consecutive amino acids toward the C-terminus of CEACAM6 that function as a sequence determinant required for addition of the glyan epitope at its upstream glycosylation site. This determinant explains the stringent specificity of the epitope for expression on CEACAM6 and, in some cancers, CEACAM5, which has significant sequence identity with CEACAM6. MAb109 that

targets this unique glycan epitope may demonstrate potential as a therapeutic for PDA, while the epitope, which is released from PDAC cells, may serve as a potential diagnostic for this disease.

INTRODUCTION

Approximately 45,000 people in the United States are diagnosed annually with pancreatic cancer. Over 90% of these cancers are classified as ductal adenocarcinoma, an aggressive cancer with a death rate nearly identical to the incidence rate; therefore, being the tenth most prevalent cancer, it is the fourth most deadly behind only lung, colon, and breast [1]. Physicians currently diagnose nearly all cases in progressive stages or after metastasis has already occurred, thereby eliminating surgical intervention, the only known curative treatment. Even with the most current detection methods, cases have usually progressed to the point where standard non-surgical therapy, which includes gemcitabine, a fluorinated nucleoside analog, with or without erlotinib, a tyrosine kinase inhibitor of the epidermal growth factor receptor (EGFR), is generally not effective [2, 3]. Although detection and treatment of many cancer types have recently improved, the survival rates for pancreatic cancer have increased minimally over the last few decades [4].

Despite neither of the current biomarkers, CECAM6 and CA19-9, being recommended as a sole screening or diagnostic marker for pancreatic cancer, many other glycans, glycoproteins, and proteins involved in glycosylation are currently under investigation as potential cancer biomarkers [4, 5]. Glycosylation is a post-translational modification involving the addition of glycans to glycoconjugates, including both

Asparagine (N-linked) and Serine/Threonine (O-linked) amino acids on proteins [6]. These glycan modifications occur in the endoplasmic reticulum (ER) and Golgi apparatus as the proteins are transported to the cell surface, endosomes, or lysosomes where they perform numerous cellular functions, which include attaching proteins to cell membranes through GPI-anchors [7], mediating cell-cell interactions [8], functioning as signaling molecules [9], and determining the folding/stability/half-life of a protein [10].

Due to the diversity of sugars, linkages, and substrates involved in glycosylation, evidence from previous studies has shown that glycan structures are dynamic and commonly altered in disease processes [11]. Glycosylation is not encoded at the gene level, therefore the synthesis of various glycan structures by the cell is very complex and depends not only on the abundance of glycosyltransferases, glycosylhydrolases, and sugar donors/acceptors, but also on their distribution within cellular compartments like the ER and Golgi. Because gene expression and cell signaling are altered in cancer cells, the glycan structures in these cells are also affected. As a result, altered glycosylation is a hallmark of oncogenic transformation. Examples of glycosylation changes associated with cancer include increased GnT-V expression and β1,6 branching of N-linked glycans [12, 13], expression of the Tn-antigen due to COSMC mutation [14], overexpression of Sialyl-Lewis-x/Sialyl-Lewis-y structures [15], loss of GPI anchors [16], and increased incorporation of Neu5Gc into sialic acid glycoconjugates [17].

Currently, nine out of ten of the most commonly used antibodies in clinical oncology as disease makers bind to glycans or glycoproteins [18]. Because of the prevalence of glycosylation changes in cancer and advances in technologies to determine glycan

structures, a focus on the identification and characterization of cancer-specific glycans is now important for the development of cell surface targets for the early detection and treatment of cancer. Glycans and glycoproteins are useful targets because changes in glycosylation often occur in the early stages of cancer progression and also provide a cell-surface molecule that could be specifically targeted in treatments, including those based on monoclonal antibodies. Several monoclonal antibodies that appeared to specifically bind glycans of cancer cells and tissues, but not those expressed in non-diseased tissues have been reported.

In this study, we describe a monoclonal antibody, MAb109, that binds to a unique, N-linked glycoepitope expressed in pancreatic cancer over non-diseased normals. Based on the screening of over 150 tissue sections from pancreatic adenocarcinoma, pancreatitis, and non-diseased pancreas, MAb109 demonstrates a sensitivity of 98% and a sensitivity of 82%. This glycoepitope is expressed on CEACAM6 and comprises an additional HexNAc residue attached to the β-linked mannose of an N-linked trisaccharide core, a glycan structure which has not been described in humans. Addition of this epitope is protein-specific, requiring amino acid recognition sequences in the C-terminus. In other cancers, the glycoepitope is expressed on CEACAM5, which shares some sequence identity with CEACAM6, and it is possible that the epitope is expressed on other proteins that are immunoglobulin super-family members. Elucidation of this N-linked glycan structure, along with the determinants required for its addition, may lead to the development of new diagnostics, and therapeutics for pancreatic adenocarcinoma and other cancers.

METHODS

Immunoblotting— BxPC3, HEK-293, HEK-Lec1, CHO wild-type, CHO-Lec8, and CHO-Lec1 cells were obtained from American Type Culture Collection (ATCC). Cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and maintained in a 37^{0C} incubator with 5% CO_2. Total cell lysates were prepared by lysing cells with RIPA buffer, and 20μg of total cell lysates were resolved on NuPAGE 4-12% Bis-Tris gels and transferred to a PVDF membrane using the Novex iBlot system. Membranes were blocked overnight at 4^{oC} in 5% non-fat milk for MAb109 and anti-Fc immunoblotting and in 3% BSA for lectin blotting with Vector Labs E-PHA. MAb109, anti-Fc, and E-PHA were used a dilution factors of 1:500, 1:2500, and 1:5000, respectively. Santa Cruz goat anti-mouse secondary antibody was used at a dilution of 1:5000 with the MAb109 and anti-Fc antibodies, and streptavidin-HRP was used at a dilution of 1:5000 with the E-PHA lecting. Membranes were incubated with PerkinElmer Western Lightening Plus-ECL according to the manufacturer's instructions before exposure.

C6f1 fragment— A 570 nucleotide fragment of CEACAM6, designated as C6f1, was PCR-amplified using a forward primer containing an EcoRI restriction site and a reverse primer containing a BglII restriction site. The PCR products were run on a 1% agarose gel to check for the product size, and the band of desired size was excised and purified using a Qiagen gel extraction kit according to the manufactures' instructions. The PCR product was then restriction enzyme digested overnight at 37^{oC} in a 20uL reaction containing 1uL of both EcoRI and BglII enzymes from Promega. The products were once again run on a 1% agarose gel to verify product size, excised from the gel, and

purified using a Qiagen gel extraction kit. Ligation of the fragment into the pFUSE vector, which had been previously treated with EcoRI and BglII, was performed using NEB T4 DNA ligase according to the manufactures' instructions. The resulting plasmid was transformed into Top 10F competent E. Coli and plated on an agar plate containing Zeocin. Single colonies were selected and expanded in 5mL cultures, and the plasmid was purified using a Qiagen mini-prep kit. DNA sequence was verified both by size on an agarose gel and by nucleotide sequencing.

Overexpression of C6f1 fragment— HEK-293 cells were transiently transfected in a 6-well dish with the C6f1-pFUSE construct using Lipofectamine 2000 according to the manufactures' protocol. Serum-free media containing the transfection reagent was replaced with DMEM containing 10% FBS after four hours, and the cells were allowed to grow overnight. The following day, cells containing C6f1-pFUSE plasmid were selected for by pressuring with Zeocin at 400ug/mL. Cells were re-fed and antibiotic added every 2-3 days for two weeks. After two weeks, the stably transfected cells were transferred to 10cm dishes and allowed to expand to 70% confluency before adding new media. The cells were allowed to grow 48 hours before the media containing secreted C6f1 was collected, and protease inhibitor cocktail was added. The media was stored at -20^{oC} until purification of C6f1 could be performed.

Purification of C6f1 fragment— Purification of C6f1 was performed using a Biologic DuoFlow Chromatography System. A GE Healthcare HiTrap HP Protein A column in combination with 20mM sodium phosphate pH 7.0 binding/wash buffer (buffer A) and

100mM citric acid pH 3.0 buffer (buffer B) were used to purify C6f1 from supernatant. The protocol for C6f1 purification with protein A column is listed below in Table 1. Fractions were taken every 5mLs up until elution with buffer B after which they were collected every 2.5mL until completion of the protocol.

Step Number	Buffer or Sample	Percentage or Gradient	Flow rate (mL/min)	Total step time (mins)
1	Buffer A	100%	5	50
2	Load/inject sample	100%	5	150
3	Buffer A	100%	5	50
4	Buffer B	100%	5	25
5	Buffer A	100%	5	50

Table 2.1. PROTOCOL FOR C6F1 PURIFICATION WITH PROTEIN A COLUMN

Twenty microliters of each fraction from Protein A column was run on polyacrylamide gel, transferred to PVDF, and blotted using MAb109 to identify fractions containing MAb109-reactive C6f1. The reactive fractions were then pooled and concentrated using an Amicon Ultra Ultracel 30KDa centrifugal filter before being brought up to a volume of 30mLs with 25mM MES pH6.5. The sample was then loaded onto the Biologic DuoFlow Chromatography system and further purified/concentrated using a GE Healthcare HiTrap Qff column with 25mM MES pH6.5 (buffer A) and 25mM MES pH6.5 with 1M NaCl (buffer B). The protocol for C6f1 purification/ concentration using the Qff column is listed below in Table 2. Fractions were taken every 5mL up until elution with buffer B after which they were collected every 2mL until completion of the protocol.

Step Number	Buffer or Sample	Percentage or Gradient	Flow rate (mL/min)	Total step time (mins)
1	Buffer A	100%	1	10
2	Load/inject sample	100%	1	30
3	Buffer A	100%	1	10
4	Linear Gradient (A-B)	Gradient	1	20
5	Buffer A	100%	1	20

Table 2.2. PROTOCOL FOR C6F1 PURIFICATION/CONCENTRATION USING QFF COLUMN

Twenty microliters of each fraction from Qff column was run on polyacrylamide gel, transferred to PVDF, and blotted using MAb109 to identify fractions containing MAb109-reactive C6f1. The reactive fractions were then pooled together and concentrated using an Amicon Ultracel 30KDa centrifugal filter and buffer exchanged into MilliQ water. Once the volume had been concentrated below 1mL, the purified C6f1 was placed in a microcentrifuge tube and completely dried in a speed-vac. The final sample was resuspended in a desired volume of milliQ and stored at -20^{oC} until being used.

Glycosidase treatment of C6f1— Approximately 1μg of purified C6f1 was left untreated or treated with combinations of neuraminidase, β-galactosidase, β-N-acetylglucosaminidase, α1-2,3 mannosidase, and α1-6 mannosidase obtained from New England Biolabs (NEB) and used according to the manufacturer's protocol. PNGaseF and EndoF enzymes were also purchased from NEB and used according to the manufacturer's protocol. Laemmli buffer containing β-mercaptoethanol was added to samples before boiling and resolving on 4-12% SDS gels. The gels were transferred to PVDF, blocked, and blotted with MAb109 and anti-Fc antibodies.

Milligram quantity digestions of C6f1 with EndoF and PNGaseF were performed using recombinant enzymes supplied by Dr. Kelley Moremen's Lab. The enzymes were used at 1μg for every 10μg of C6f1. Digestions were incubated at 37^{oC} for 48 hours to insure complete digestion.

Trypsin/Chymotrypsin digestion of C6f1— Approximately 1mg of purified C6f1 was resuspended in 200ul of 0.1M Tris-HCL pH 8.2 containing 10mM $CaCl_2$ in a microcentrifuge tube. The tube was heated to 100^{oC} for 5 minutes and then cooled to room temperature. 25uL of each freshly prepared trypsin and chymotrypsin (2mg/mL) were added to the reaction and incubated at 37^{oC} overnight. The tube was spun down in a microcentrifuge at max speed at 4^{oC} for 10 minutes. Supernatant was carefully removed to another tube before adding 200uL of water to the remaining sample. The tube was mixed well and then centrifuged again at max speed at 4 ^{OC} for 10 minutes. 200uL of the supernatant was collected and combined with the previous sample in the same tube. The sample was dried down completely in a speed-vac before being fractionated and used in subsequent experiments.

Inhibition assays— 100ng of purified C6f1 was spotted in triplicate for each reaction on PVDF membrane using a vacuum manifold. The membrane was blocked in 5% milk solution with TBS-T overnight at 4^{oC}. The stock MAb109 (0.1ug/uL) was diluted 1:3000 to a final concentration of 0.03ng/uL in blocking solution before each competitor was added and incubated for 30 minutes at room temperature. Each competitor had been prepared starting with 100ng of either C6f1 or RNaseB glycoprotein or were fractions

collected from reverse-phase separation of C6f1 glycopeptides. C6f1 and RNaseB glycoproteins were treated with various glycosidases and proteases, including PNGaseF, trypsin, and chymotrypsin, according to the manufacturer's instructions in order to produce glycans, glycopeptides, and peptides. Glycans and peptides were separated using a Waters Sep-Pak C_{18} column by first equilibrating the column with three volumes of 100% acetonitrile. The column was then washed with three volumes of 5% acetic acid solution before loading the sample, which had been acidified to final concentration of 0.1% trifluoroacetic acid (TFA). The flow through fraction containing glycans was dried down in a speed-vac, while the peptides were eluted with one column volume of 100% acetic acid before also being dried down in a speed-vac. The antibody/competitor solution was then incubated with the C6f1 spotted strip for 1 hour at room temperature. After 1 hour, the strips were washed three times with TBS-T for 5 minutes each wash. The secondary antibody was Santa Cruz anti-mouse IgG HRP diluted 1:5000 and incubated at room temperature for 30 minutes. Three TBS-T washes were repeated before adding Perkin-Elmer Western Lightening Plus-ECL according to the manufacture's instructions. Blots were exposed to x-ray film for 10 minutes before developing. After developing, the film was scanned, and densitometry was performed using ImageJ software.

Fractionation of C6f1 glycopeptides using reverse phase chromatography- Dried tryptic peptides were resuspended with 50mL of buffer A and then separated by offline reverse phase liquid chromatography. Solvent A (0.1% TFA) and solvent B (80% acetonitrile, 0.085% TFA) were used to develop a linear gradient consisting of 5 min at 95% solvent

A, 60 min gradient at variable slope to 95% solvent B, 3 min at 95% solvent B, 5 min to 95% solvent A, and 16 min at 100% solvent A. Fractions were collected every 5 min, and pooled based upon the chromatogram. Pooled fractions were dried down in a SpeedVac. Dried glycopeptides fractions of interest were resuspended with 50mL of Solvent A and separated by reverse phase using the same gradient. Fractions were collected every 5 minutes and tested for the presence of the MAb109 epitope using the competitive inhibition assay described previously.

Release of N-glycans and 18-O labeling of N-glycosylations sites— Glycopeptides were reconstituted with 45ul of 18O-water (97% Cambridge Isotope Laboratories Inc.) and 5ul of 0.1 M $NaPO_4$ buffer, pH7.5. PNGaseF (3ul) was added an incubated at 37 ^{OC} overnight. Reaction was stopped by boiling for 5 minutes. Solution was dried by speed-vac concentrator. Released N-glycans were separated by passing through a C_{18} column. Unbound fractions and wash fractions were lyophilized for permethylation of N-glycans. Bound materials were eluted with 20% and 40% 2-PrOH in %5 AcOH and dried by speed-vac. To delabel the peptide C-terminal 18O incorporation, peptides were reconstituted in 40mM ammonium acetate buffer, pH6.5 and re-treated with 1ug of chymotrypsin overnight. Peptides were dried by speed-vac concentrator and reconstituted with 25uL of 0.1% formic acid.

Peptide analysis— Sample was loaded off-line onto a nanospray tapered capillary column/emitter self-packed with C18 reverse-phase resin in a nitrogen pressure bomb for 5 minutes at 1000psi (~3uL) and then separated via 160-minute linear gradient of

increasing mobile phase B. LC-MSMS was performed with a program by which the 5 most intense peaks on each full mass spectrum were captured to be subjected to fragmentation (Top6) MS data was analyzed using Sequest requiring fully tryptic/ chymotryptic peptides and allowing for dynamic modification of Cys +57 (alkylation with iodoacetamide), Met +16 (oxidation), and Asn +3 (site of N-linked glycosylation) with Proteome Discoverer software (ThermoScientific).

N-linked Glycan release and analysis— N-linked glycans were released from tryptic/chymotryptic digests of C6f1 glycoprotein by enzymatic digestion with PNGaseF (Aoki 2007). C6f1 glycoprotein was resupended in 200 ml of trypsin buffer (0.1 M Tris-HCl, pH 8.2, containing 1 mM $CaCl_2$). After boiling for 5 min and cooling to room temperature, 25 ml of trypsin solution (2 mg/ml in trypsin buffer) and 25 ml of chymotrypsin solution (2 mg/ml in trypsin buffer) were added. Digestion was allowed to proceed for 18 h at 37 °C before the mixture was boiled for 5 min. The reaction mixture was dried by vacuum centrifugation. The dried peptide/glycopeptide mixture was resuspended in 250 ml of 5% acetic acid (v/v) and loaded onto a Sep-Pak C18 cartridge column. The cartridge was washed with 10 column volumes of 5% acetic acid. Glycopeptides were eluted stepwise, first with 1 volumes of 20% isopropyl alcohol in 5% acetic acid and then with 1 volumes of 40% isopropyl alcohol in 5% acetic acid. The 20 and 40% isopropyl alcohol steps were pooled and evaporated to dryness. Dried glycopeptides were resuspended in 50 ml of 50 mM sodium phosphate buffer, pH 7.5 for digestion with PNGaseF. Following PNGaseF digestion for 18 h at 37 °C, released N-linked glycans were separated from peptide and enzyme by passage through a Sep-

Pak C18 cartridge. The digestion mixture was adjusted to 5% acetic acid and loaded onto the Sep-Pak. The column run-through and an additional wash with 3 column volumes of 5% acetic acid, containing released glycans, were collected together and evaporated to dryness.

Released N-linked glycans were permethylated, dissolved in 1 mM sodium hydroxide in methanol/water (1:1) for infusion into the mass spectrometer. Nanospray ionization mass spectrometry was performed on permethylated N-glycans. For MS of permethylated N-glycans in positive ion mode, permethylated N-glycans were dissolved in 50 ml of 1 mM sodium hydroxide in methanol/water (1:1) for infusion into a linear ion trap mass spectrometer (Orbi-LTQ; Thermo Fisher Scientific, Waltham, MA) using a nanoelectrospray source at a syringe flow rate of 0.40 µl/min and capillary temperature set to 210°C. The instrument was tuned with N-glycan standard obtained from bovine pancreatic ribonuclease B for positive ion mode. For fragmentation by collision-induced dissociation (CID) in MS/MS and MS^n, a normalized collision energy of 30% to 35% was used.

Detection of individual N-glycans was accomplished using the total ion mapping (TIM) and neutral loss scan (NL scan) functionality of the Xcalibur software package version 2.0 (Thermo Fisher Scientific). For TIM, the m/z range from 600 to 2000 was automatically scanned in successive 2.8 mass unit windows with a window-to-window overlap of 0.8 mass units, which allowed the naturally occurring isotopes of each N-glycan species to be summed into a single response, thereby increasing detection sensitivity. Most N-glycan components were identified as singly, doubly, and triply charged, sodiated species (M+Na) in positive mode. Preliminary analysis demonstrated

that the characteristic fragment ions in CID MS/MS scans of C6f1 N-glycans correspond to the neutral loss of HexNAc or sialic acid moiety. Therefore, an MS workflow was defined for NL scans in which the characteristic ions detected by full MS was subjected to CID fragmentation. If an MS/MS profile contained an ion with m/z equivalent to loss of the non-reducing end of HexNAx moiety, MS^n fragmentation was initiated. Following this data-dependent acquisition, the workflow returned to the full MS, excluded the parent ion just fragmented, and chose the peak of the interest ion for the same MS/MS and MS^n analysis.

NMR analysis— NMR samples were prepared by dissolving lyophilized material in 99.96% D2O. Spectra were acquired on Agilent DD2 600 and 900 MHz spectrometers using standard pulse sequences. Data were collected at 10C,25C and 38C in order to move the HDO signal away from signals of interest. One-dimensional proton spectra were acquired using a presaturation pulse to suppress the residual HDO signal. Two-dimensional TOCSY (total correlation spectroscopy) and ROESY (rotating frame NOE) data sets were collected with a spectral width of 11.6 ppm, mixing times of 80 ms and 200 ms, and 32 and 128 transients per increment, respectively. The TOCSY was collected as a 2095 X 240 complex points data set, and the ROESY as a 1047 X 320 data set. Data was then processed with MestreNova software (MestreLab, Inc.).

Expression of C6f1 in HEK Lec1 suspension cells— Cells were propagated and maintained in a 50:50 mix of Ex-Cell/Freestyle media until time of transfection. At the time of transfection, the cells are collected by centrifugation and resuspended in

FreestyleTM 293 medium alone at a cell density of 2.5 x 10^6 cells/mL. Cells were then transfected using DNA and polyethylenimine (PEI) at final concentrations of 2.5ug/mL and 0.5 ug/mL, respectively, in FreestyleTM 293 medium. Cells were maintained in FreestyleTM 293 medium for 24 hours post-transfection. After 24 hours, the cultures were diluted 1:1 with Ex-Cell or ESF medium with valproic acid(VPA) added to a final concentration of 2.2 mM. Cultures were then maintained for recombinant glycoprotein production for five days on a platform shaker in a CO_2 incubator. After five days, cultures were centrifuged to separate the cells from the suspension media containing the desired C6f1 glycoprotein. Additionally, cell culture media was passed over a 0.42um vacuum filter before purifying using a BioRad chromatography system with a Protein A column.

Cloning and expression of CEACAM6 fragments/mutants— PCR cloning of fragments of C6f1 was performed using HotStarTaq DNA Polymerase by Qiagen according to the manufacturer's instructions using C6f1-pFUSE as the template. The forward and reverse PCR primers were designed with EcoRI and BlgII restriction sites, respectively, in order to facilitate cloning into pFUSE vector. Primer sequences are as follows: Ig1F: GAATTCAAAGCCCTCCATCTCCAGC, Ig1R: AGATCTATTCAGGGTGACTGGGTC, Ig2F: GAATTCAGATGGCCCCACCATTTGG, Ig2R: AGATCTGGTGACTGTGGTCCTATT, peptide3F: GAATTCACTGCAGCTGTCCAATGGC, and peptide3R: AGATCTTTTGACGCTGAGTAGAGT. PCR products were restriction enzyme digested with EcoRI and BglII transformed using Promega restriction enzymes,

ethanol/chloroform-precipitated, and ligated into a prepared pFUSE vector using NEB T4 DNA ligase according to the manufacturers instructions. Top10F competent cells were transformed and plated on agar plates containing Zeocin. Several clones were selected, grown, and plasmid prepped before inserts being verified using restriction enzyme digests with EcoRI and BglII as well as DNA sequencing.

Site directed mutagenesis of C6f1 and C8f1 was performed using a Stratagene QuikChange site-directed mutagenesis kit according to the manufacturer's protocol. The primers for both the C6f1 ^{300}QAH^{302} to ^{300}HTT^{302} and C8f1 ^{300}HTT^{302} to ^{300}QAH^{302} were C6f1 ^{300}HTT^{302} forward: GCGGATCCTATATGTGCCACACCACTAACTCAGCCACTGGCCTC and reverse: GAGGCCAGTGGCTGAGTTAGTGGTGTGGCACATATAGGATCCGC. C8f1 ^{300}QAH^{302} forward: GGATCCTATGCCTGCCAAGCCCATAACTCAGCCACTGGC and reverse: GCCAGTGGCTGAGTTATGGGCTTGGCAGGCATAGGATCC. PCR products were transformed using XL-1 blue competent cells and plated on agar plates containing Zeocin. Several clones were selected, grown, and plasmid prepped before inserts being verified using restriction enzyme digests with EcoRI and BglII as well as DNA sequencing.

RESULTS

The MAb109 epitope is an N-linked glycan expressed on CEACAM6 in pancreatic adenocarcinoma (PDAC) cultured cells. Previous immunohistochemical (IHC) staining by Krueger had shown that MAb109 reacts with pancreatic cancer tissues. In order to

gain insight into the prevalence of the MAb109 epitope in PDAC, we performed SDS-PAGE and immunoblots on total cell lysates from BxPC3 cells, a PDAC cell line (Fig. 2.1). The left side of panel 2.1A shows a broad, MAb109-reactive band with an observed molecular weight centralized at approximately 85 kDa. Similar results had been observed in previously unpublished studies by Gelber in which total cell lysates from small cell lung cancer (SCLC) were subjected to SDS-PAGE and immunoblotted with MAb109. In these studies, MAb109 reactivity was shown to overlap with that of a glycoprotein, Carcinoembryonic Antigen 6 (CEACAM6). Furthermore, MAb109 reactivity was abolished when the small cell lung cancer (SCLC) total cell lysate was treated with PNGaseF, a glycosidase that removes N-linked glycans from human proteins. Since CEACAM6 is predicted to contain 12 N-linked glycosylation sites and is overexpressed in pancreatic cancer, we hypothesized that the MAb109 was binding to an N-linked glycan or glycans of CEACAM6 in pancreatic cancer.

To test the hypothesis that the MAb109 epitope involved an N-linked glycan on CEACAM6, we treated the BxPC3 total cell lysate with PNGaseF. BxPC3 total cell lysate treated with PNGaseF no longer demonstrated MAb109 reactivity (Fig. 2.1A). This observation strongly suggested the epitope was itself an N-linked glycan or that an N-linked glycan was in some way responsible for antibody reactivity. Additionally, immunoblotting with a commercially available anti-CEACAM6 polypeptide antibody verified that the protein was present after PNGaseF treatment (Fig. 2.1B). The molecular weight of CEACAM6 shifted from approximately 85 kDa to 39 kDa after removal of the N-linked glycans, which supports the protein being glycosylated (Fig. 2.1B). Similar results were obtained with Capan-1 cells, another PDAC cell line (data

not shown). Taken together, these results suggested that the MAb109 epitope in both these cell lines was likely an N-linked glycan or glycans expressed on CEACAM6.

CEACAM6 is a glycophosphatidylinositol-(GPI) anchored glycoprotein belonging to the immunoglobulin (Ig) superfamily and contains 12 sequons for N-linked glycosylation. CEACAM6 is a member of the Carcinoembryonic Antigen Cell Adhesion Molecule (CEACAM) family of proteins, which mediate cell adhesion and are involved in the maintenance of cell polarity and tissue architecture. The CEACAMs contain Ig-domains and several predicted N-linked glycosylation sequons. CEACAM6 has a total of three immunoglobulin domains, one N-terminal variable-type (V) domain and two C-terminal constant (C1 and C2) domains, and also has twelve predicted N-linked sequons. Panel A of Figure 2.2 shows a graphical representation of the CEACAM6 with immunoglobulin domains (V, C1, and C2) depicted as loops held together by disulfide bridges and N-linked glycosylation sites as ball-and-stick extensions. The GPI-anchor is depicted graphically as a triangle on the C-terminus of CEACAM6.

To show the distribution of N-linked glycosylation sites, the amino acid sequences for the N-terminal variable domain (V), CEACAM6 fragment 1 (C6f1) consisting of the two C-type domains (C1 and C2), and the GPI-anchor are shown in relation to predicted N-linked sites (Fig. 2.2B). Predicted N-linked sites following the N-X-S/T motif were shown as enclosed boxes.

CEACAM6 fragment 1 (C6f1) transiently transfected in HEK-293 cells expresses the MAb109 epitope that is PNGaseF sensitive. MAb109 reacted with pancreatic cancer

cell lines and tissues. To test whether the MAb109 epitope could be expressed *in vitro* in a non-cancerous, non-pancreas cell types, we constructed a secreted fragment of CEACAM6 by inserting a 190 amino acid C-terminal fragment of CEACAM6 without its GPI anchor into the pFUSE vector. This vector encodes for an N-terminus IL-2 signal sequence, which targets glycoprotein to the secretory pathway, as well as a C-terminus Fc-fusion tag (Fig. 2.3A). The resultant CEACAM6 fragment 1, denoted C6f1, construct encodes for the constant Ig-domains of CEACAM6 and reduces the number of predicted N-linked sequons from twelve to nine. BxPC3 transfected with C6f1 expressed the construct at low levels (data not shown).

To increase the expression of C6f1 and potentially the MAb109, we transfected HEK-293 cells, a cell line commonly used to express recombinant protein, with the C6f1 construct. Supernatant was collected from C6f1 transiently transfected HEK-293 cells and subjected to SDS-PAGE followed by immunoblotting with MAb109. The data revealed in the observance of a 85kDa MAb109-reactive band (Fig 2.3A), which was not present in either supernatants or total cell lysates before or after transfection of the pFUSE vector only control. These results show that HEK-293 cells contain the biosynthetic machinery required for MAb109 epitope expression, and that the cells express this epitope only when transfected with the C6f1 amino acid sequence. The levels of the epitope were also much less than in BxP3 cells expressing C6f1 (data not shown).

Additionally, C6f1 was treated with PNGaseF to determine whether the MAb109 epitope synthesized by HEK-293 cells was expressed on an N-linked glycan similarly to PDAC cells. The left panel of Figure 2.3C shows that PNGaseF treatment of the

recombinantly expressed C6f1 totally abolished MAb109 reactivity. Additionally, the right panel of Figure 2.3C shows anti-Fc antibody immunoblotting confirming the presence of the protein after glycosidase treatment. The size of C6f1 shifted from 85 to 50 kDa, confirming that C6f1 was N-linked glycosylated. Since PNGaseF abolished MAb109 reactivity, we next sought to use cell lines and glycosidases to alter the structures of N-linked glycans and possibly affect MAb109 binding. This would allow the identification of structures or specific sugars required for MAb109 binding.

An overview of N-linked biosynthesis is shown in Figure 2.4. High-mannose, hybrid, and complex glycans are all synthesized from a common $Glc_3Man_9GlcNAc_2$ precursor that is added by oligosaccharyltransferase (OST) to newly synthesized proteins in the endoplasmic reticulum (ER). After addition of $Glc_3Man_9GlcNAc_2$, the glycan undergoes processing as terminal glucose residues are cleaved. Once correctly folded, the glycoprotein is trafficked through the ER, and the glycan undergoes successive mannose trimming by α-mannosidases, which ultimately yields a key intermediate, $Man_5GlcNAc_2$. Prior to this intermediate, the glycans are high-mannose glycans. *N*-acetylglucosaminyltransferase-I (GnT-I) can add a GlcNAc residue to the C-2 of the α1-3 mannose in the core of Man_5Glc_2, which allows the biosynthesis of hybrid and complex glycans. Lec1 cells, cell lines that have GnT-I mutations, only express high-mannose glycans because they cannot synthesize the immediate precursor substrate required for hybrid and complex glycans.

Whether or not the subsequent GnT-I product, $Man_5GlcNAc_3$ is acted on by α-mannosidase II will determine whether the glycan develops into a complex or hybrid glycan, respectively. If α-mannosidase II does act on the glycan, then GnT-II adds a β1-

2 linked GlcNAc to the α1-6 linked mannose which creates a biantennary glycan, a structure that can be acted on by subsequent GlcNAc transferases to increase branching. The GlcNAc transferase and the linkage of the sugar that it adds are designated in Figure 2.4. After addition of branching GlcNAc residues, the sugars can be extended with galactose, sialic acid, and poly-*N*-acetyllactosamine.

MAb109 reacts with C6f1 recombinantly expressed in CHO wild-type and glycosylation deficient mutant cell lines. After the expression of the MAb109 epitope in HEK-293 cells, we anticipated that other cell lines could be used to express C6f1 and aid in the characterization of the MAb109 epitope. Figure 2.5 shows SDS-PAGE and MAb109 immunoblotting of supernatants collected from CHO wild-type, Lec8 (galactose transporter deficient), and Lec1 (GnT-I deficient) cell lines that had been transiently transfected with either the pFUSE or C6f1 construct. The MAb109-reactive C6f1 secreted from CHO wild-type cells had the same 85kDa molecular weight as did the C6f1 from HEK cells, yet the C6f1 produced in Lec8 and Lec1 cells shifted to lower molecular weights, approximately 75kDa and 62kDa, respectively (Fig. 2.5A). MAb109 eptiotpe expression might have been affected since the mutation in Lec8 cells results in a lack of galactose and sialic acid addition to N-linked glycans, while the mutation in Lec1 cells results in the synthesis of only high-mannose glycans, but MAb109 reactivity was observed in both these cell lines. Therefore, these cell lines contain the biosynthetic machinery to express the MAb109 epitope, and the binding epitope does not require a hybrid or complex N-linked glycan since Lec1 cells can synthesize only high-mannose glycans. The epitope may be attached to one of the sugars of the

trisaccharide core since this structural feature is common to all three types of N-linked glycans.

Additionally, the anti-Fc reactivity shows that all constructs were expressed, and the pFUSE only transfections demonstrated that protein was expressed but does not received the MAb109 epitope (Fig. 2.5B). The C6f1 sequence was required for addition of the MAb109 epitope since expression was not observed with the pFUSE plasmid control. Observing MAb109 reactivity in glycosylation mutant Lec8 and Lec1 cells provided evidence that these cells contain the biosynthetic machinery to synthesize the MAb109 epitope when the C6f1 fragment was expressed even though the epitope was not expressed endogenously.

MAb109 reacts with a modification found on the trisaccharide core of an N-linked glycan of C6f1. To characterize the N-linked glycans on which the MAb109 glycoepitope is found by another manner besides glycosylation-deficient cell lines, purified C6f1 from HEK-293 cells was digested using a combination of glycosidases that hydrolyze the glycosidic bonds of sialic acid (neuraminidase), β-galactose (β-galactosidase), β-N-acetylglucosamine (β-N-acetylglucosaminidase), and α-mannose (mannosidases) from N-linked glycans. The top panel of Figure 2.6 shows the glycosidases that were included in each C6f1 reaction.

After overnight digestion of C6f1 with the glycosidases, the reaction products were subjected to SDS-PAGE and immunoblotting with MAb109. MAb109 immunoblotting of glycosidase treated C6f1, shown in the lower panel of Figure 2.5, revealed that all enzymatic reactions remained positive for MAb109 reactivity despite shifts to

progressively lower molecular weights with successive glycosidase treatment. Additionally, MAb109 reactivity remained constant for all reactions after the removal of terminal glycans. Glycosidases were functional since mass shifts of the glycoprotein were observed after each sequential enzyme digestion. Taken together with MAb109 expression in CHO mutant cell lines, these findings further support the findings in Figure 2.5 that the epitope is attached to one of the sugars of the trisaccharide core. This structural feature is common to N-linked glycans and not removed by glycosidase treatment.

Milligram quantities of recombinant C6f1 can be purified using transiently transfected HEK-293 cells. In order to obtain sufficient amounts of Fc-tagged glycoprotein expressing the MAb109 epitope for biochemical and structural studies, HEK-293 cells were transiently transfected with the C6f1 construct, and the expressed glycoprotein was purified. Purification was performed on transfected cell supernatants using immobilized protein A by fast protein liquid chromatography (FPLC). The chromatography trace in Figure 2.7A plotting absorbance at 220 nm and 280nm shows that the C6f1 glycoprotein binds to the column at a neutral pH with 20mM sodium phosphate buffer pH 7.0 (buffer A) and is eluted under acidic conditions with 100mM citric acid pH 3.0 (buffer B). Elution of C6f1 was measured by absorbance at 220 nm and 280 nm and corresponded to an increased absorbance peak with fractions 31 to 36. To test for the expression of the MAb109 epitope on the purified C6f1, eluted fractions were subjected to SDS-PAGE followed by MAb109 immunoblotting. Immunoblots showed a single, 85 kDa MAb109-reactive band in fractions 31 to 36 (Fig. 2.7B).

MAb109-reactive fractions were pooled, concentrated, and buffer-exchanged to water before a sample was subjected to SDS-PAGE and silver stained in order to check the purity and concentration of the C6f1. An 85 kDa band was observed on silver stain, and the amount of recombinant C6f1 was calculated by comparing protein levels against bovine serum albumin (BSA) standards on the same gel (Fig. 2.7C). The yield of C6f1 stably expressed in HEK-293 cells was approximately 1μg/mL for the supernatant collected from these cells.

Glycans and glycopeptides derived from C6f1 inhibit MAb109 binding. After previously showing that various cell types expressed MAb109-reactive C6f1 and that the epitope may be expressed on the trisaccharide core, we sought to provide further evidence that MAb109 binding requires a N-linked glycan or glycans specifically expressed on C6f1. We hypothesized that if the MAb109 epitope involved an N-linked glycan or glycans specific to C6f1 then glycans and glycopeptides prepared from C6f1 should inhibit MAb109 binding. To test our hypothesis, we developed a competitive inhibition assay, shown in Figure 2.8, to test whether various peptides, glycans, or glycopeptides prepared and isolated from C6f1 and a control glycoprotein, RNaseB, could inhibit MAb109 binding.

Competitors for MAb109 binding were each derived from 0.1 ug of C6f1 that was treated with various glycosidases and proteases, including PNGaseF, trypsin, and chymotrypsin, in order to produce glycans, glycopeptides, and peptides. To demonstrate that MAb109 inhibition was specific to C6f1 glycans and glycopeptides, RNaseB, a glycoprotein known to be N-linked glycosylated but not to receive the

MAb109 epitope, was used as a control and subjected to the same enzymatic treatments as C6f1. The left panel of Figure 2.9A identifies the competitor used for each numbered, C6f1 spotted PVDF strip. The center panel shows MAb109 reactivity for each competitor tested. Inhibition of MAb109 reactivity was observed only for C6f1 glycopeptides and glycoprotein, but not for C6f1 lacking N-linked glycans or glycans derived from RNaseB. The upper right panel shows immunoblots for anti-Fc reactivity, which was used to normalize the amount of C6f1 on the strip, demonstrating that the glycoprotein was present.

Inhibition of MAb109 binding +/- competitor to C6f1, shown in Fig. 2.9B, was quantified using densitometry measurements, comparing the ratio of MAb109 signal intensity to anti-Fc signal intensity for the dot blots shown in panel A. The ratio for of MAb109 signal without competitor to its anti-Fc signal was set to 100% of MAb109 signal, to which all other competitors were compared. The process of measuring the MAb109 signal for each competitor, standardizing to the anti-Fc signal for the amount of C6f1 spotted, and comparing to the MAb109 signal without competitor was performed for each competitor. Error bars reflect error between the ratio of MAb109 signal to anti-Fc signal amongst triplicate spots.

Intact C6f1 and C6f1 glycans almost completely inhibited MAb109 reactivity in the competition assay while, C6f1 glycopeptides slightly inhibited. Otherwise, no statistically significant inhibition was observed with deglycosylated C6f1 protein or peptides or with any competitors derived from RNaseB. C6f1 derived N-linked glycans were common to intact C6f1, C6f1 glycopeptides, and C6f1 glycans, which provides additional evidence that the MAb109 epitope is an N-linked glycan expressed on C6f1.

C6f1 glycopeptide inhibition of MAb109 binding is concentration dependent. A competitive inhibition assay using C6f1 glycopeptides resulted in parital MAb109 inhibition. To test whether C6f1 glycopeptides could completely inhibit MAb109 like intact C6f1, we incubated MAb109 with increasing amounts of C6f1 glycopeptides, as measured by the amount of C6f1 treated with trypsin and chymotrypsin. The left panel of Figure 2.10A shows that increasing the amount of C6f1 glycopeptides decreases the amount of MAb109 reactivity. Complete inhibition of the MAb109 signal results when C6f1 was in approximately 500 fold excess to MAb109 (lane 5). The right panel shows anti-Fc reactivity which were used to normalize the amount of C6f1 spotted on each strip.

Inhibition of MAb109 binding by C6f1 glycopeptides was quantified using densitometry measurements in the same manner that was performed on competition inhibition assays for competitors isolated from C6f1 and RNase B (Figure 2.9) and is graphed in Figure 2.10B. No statistically significant inhibition was observed for C6f1 glycoproteins incubated in 100-fold excess of MAb109; however, the MAb109 signal was less than decreased when an excess of competitor to MAb109 was used at a ratio of 250:1. Complete inhibition was observed at a competitor to antibody ratio of 500:1. These findings demonstrate that the MAb109 epitope is expressed on glycoproteins specific to C6f1 and that these glycopeptides inhibit MAb109 in a concentration dependent manner. Additionally, inibition occcured with glycan only isolated from C6f1 peptides, not RNaseB, demonstrating that MAb109 does not likely require amino acids

for binding, nor does it recognize secondary or teriary protein structure (Fig. 10C).

MAb109-reactive glycopeptides can be isolated using reverse-phase chromatography and competitive inhibition. Because we were able to show that the MAb109 epitope was expressed on C6f1 glycopeptides, we next continued to use the assay to aid in the isolation of MAb109-reactive glycopeptides(s) that could be used for mass spectrometry sequence and structure analysis. Using this technique, a milligram of recombinant C6f1 glycoprotein was treated with trypsin and chymotrypsin to produce glycopeptides before separating MAb109-reactive from non-reactive species using reverse-phase chromatography (Fig. 2.11A). C6f1 glycopeptides were loaded in 0.1% trifluoroacetic (TFA) acid and eluted with a linear gradient consisting of an increasing percentage of 80% acetonitrile, 0.085% TFA solution. The glycopeptides were monitored by absorbance at 220 nM. Fractions, shown below the chromatography trace, were collected every five minutes and dried down before testing for competitory activity.

Competitive inhibition assays revealed that only three of sixteen fractions, fractions 11, 12, and 13, inhibited MAb109 binding, as demonstrated by the complete absence of MAb109 reactivity (Fig. 2.11B). The identical dot-blot PVDF sections were stripped and immunoblotted with anti-Fc antibody as a control to show that the lack of MAb109 reactivity was not due to lack of protein. The three MAb109-reactive fractions were pooled for subsequent peptide sequencing and glycan analysis by MS.

MAb109 epitope is found on a single, 16 amino acid peptides of C6f1. To determine which N-linked site or sites carried the MAb109 epitope, MAb109-reactive glycopeptides

from fractions 11-13, which were isolated using reverse phase chromatography, were treated with PNGaseF. Peptides were separated from glycans using C18. LC-MSMS of the isolated peptides was performed with a program by which the 5 most intense peaks on each full mass spectrum were captured to be subjected to fragmentation (Top6) MS data was analyzed using Sequest requiring fully tryptic/chymotryptic peptides and allowing for dynamic modification of Cys +57 (alkylation with iodoacetamide), Met +16 (oxidation), and Asn +3 (site of N-linked glycosylation) with Proteome Discoverer software (ThermoScientific). The MS spectra of the peptides with the most abundant peaks are labeled in Figure 2.12A. The most abundant ions observed were y-ions in the 1+ charge state, although all m/z peaks for b- and y-ions of the peptide sample are shown in Figure 2.12B. The figure shows a table of m/z peaks corresponding the b- and y- ions for amino acids identified for a 16 amino acid peptide. The two left columns show the m/z values for b-ions in the 1+ and 2+ charge states while the two right columns show m/z values for y-ions in the 1+ and 2+ charge states. The values correspond to the amino acid sequence of the peptide shown in blue column and match the ^{191}LQLSNGNMTLTLLSVK^{207} sequence of both C6f1 and CEACAM6. MS Peptide analysis identified a single 16 amino acid peptide corresponding to amino acids 191 to 207 of CEACAM6 sequence, which contained a predicted N-linked glycosylation sequon shown in red. This site corresponds to site 3 of the C6f1 fragment, ^{197}NMT^{199}. Only one peptide sequence was identified for the isolated, MAb109-reactive glycopeptides, demonstrating that epitope addition is not only protein specific but also N-linked site specific.

MAb109-reactive glycan from HEK-293 cells is a biantennary glycan terminating in galactose containing an additional HexNAc attached to the β-linked mannose. To determine the structure of the MAb109-reactive glycan or glycans, N-linked glycans purified from fractions 11-13 were analyzed after permethylation using nanospray mass spectrometry (Fig. 2.13). The MS spectrum obtained by total ion monitoring showed that the most abundant MAb109-reactive glycan was a biantennary N-linked glycan terminating with galactose with an additional HexNAc, found in two different mass-to-charge states (2+ and 3+). This glycan corresponds to peaks 1138.20 (z=3) and 1706.80 (z=2) in the MS spectra. Additionally, very little sialic acid addition was observed, which is common for HEK cell N-linked glycans, while no tri- and tetra-antennary structures were observed. Neither the linkage of the additional HexNAc, nor the identity of the sugar, could be identified by the total MS spectrum; therefore, the N-linked structures containing the additional HexNAc were annotated with an additional GlcNAc or GlcNAc over the glycan in Figure 2.13.

To determine the structure of the N-linked glycan bound by MAb109, MS^n analysis was performed on the parent ion at 1706 shown in Figure 2.13, and the results are shown in Figure 2.14. The parent ion shown was a biantennary N-linked glycan terminating in galactose and had an additional HexNAc. The scheme for analyzing MS/MS spectra is shown in Figure 2.14. MS^2 shows peaks at m/z values of 938, 836, and 706 which are values that correspond to predicted mass glycans which have an additional HexNAc attached to the β-linked mannose. Successive MS^3 was performed on the most abundant 938 peak from MS^2, revealing fragments that corresponded to the HexNAc being attached to the β-linked mannose. The m/z peaks of 1112 in MS^4 and of

907 in MS^5 provide fragments with the HexNAc being attached to the β-linked mannose. MS^6, MS^7, and MS^8 are spectra showing fragmentation of the mannoses and second GlcNAc of the chitobiose core attached to the β-linked mannose, and the spectra were consistent with fragmentation expected if the HexNAc were attached to the β-linked mannose. The m/z peaks that could have distinguished the HexNAc residue as GalNAc or GlcNAc were not observed.

HexNAc residue could be attached to the β-linked mannose at either the 2 or 4 position. Since MS^n analysis of MAb109-reactive glycans from HEK cells could not distinguish the additional HexNAc residue as either GalNAc or GlcNAc, nor determine its linkage to the β-linked mannose, the MAb109 epitope structure was predicted based on the linkages the HexNAc could form with the free hydroxyl groups of the β-linked mannose. Mannose is depicted in the chair conformation in Fig. 2.15A, as this is the conformation found when incorporated into N-linked glycans. In this conformation, mannose has five free hydroxyl groups capable of forming glycosidic bonds. In an N-linked glycan, the GlcNAc from the chitobiose core is bound in the β-conformation at the 1-position, Man is bound in the α-confomation at the 3-position, and an additional Man is bound in the α-conformation at the 6 position. As a result, the only free hydroxyl groups available for the formation of additional glycosidic linkages are at the 2- and 4- positions. The additional HexNAc identified linked to the β-linked mannose by MS must be at one of these two positions.

Because of the two possible positions, α or β1-2 or α or β1-4, plus two possible sugars, GlcNAc or GalNAc, eight possible structures exist for the MAb109 epitope (Fig.

2.15B). The β1,4-linked GlcNAc is the only structure of the eight that has been previously been described. This structure is the product of GnT-III and results in bissected glycans. The other seven predicted structures have not previously descibed in humans, and likealy represent a novel, N-linked structure for the MAb109 epitope.

HEK-Lec1 cells express the MAb109 epitope. To test the whether the MAb109 epitope was the β1,4-linked GlcNAc product of GnT-III, we expressed C6f1 in HEK Lec1 cells, which lack GnT-I activity. HEK Lec1 cells, therefore, lack the enzyme activity required to synthesize the substrate necessary for GnT-III and can only produce high-mannose glycans. If the MAb109 epitope was a β1,4-linked GlcNAc, it could not be biosynthetically expressed in HEK-Lec1 cells. Shown in the left panel of Figure 2.16, the HEK Lec1 cells transfected with C6f1 expressed an approximately 63 kDa MAb109-reactive band. This band was approximately 22 kDa lower when compared to C6f1 expressed in HEK wild-type cells (Fig. 2.16). The center panel shows that anti-Fc reactivity reflective of C6f1 expression. The MAb109 and anti-Fc reactive bands in HEK-Lec1 cells were not as broad as HEK-293 cells, suggestive of reduced glycan heterogeneity between the two cell lines. Additionally, the far right panel shows showed the reactivity of *Phaseolus vulgaris* Erythroagglutinin (E-PHA), a lectin specific for the bisected glycan, with C6f1 expressed in HEK-293 cells, but E-PHA reactivity was absent in HEK-Lec1 cells. This observation was expected since HEK-293 cells can synthesize the bisected glycan, but HEK-Lec1 cells cannot. Since C6f1 has nine predicted N-linked sites, the E-PHA reactive glycan or glycans in HEK-293 cells are likely added to an N-linked site other than Asn^{197}, while the MAb109 reactive glycan

itself is not reactive. This hypothesis was supported by the expression of C6f1 in a siRNA GnT-III knockdown cell line prepared by Dr. Karen Abbott. E-PHA reactivity was decreased in siRNA GnT-III knockdown cells while MAb109 reactivity was unaffected when compared to a scrambled siRNA control cell line (data not shown). The presence of MAb109 reactivity and absence of E-PHA reactivity for C6f1 expressed in Lec1 cells, taken together, suggest that the MAb109 epitope is not the bisected β1,4-linked GlcNAc.

MAb109 epitope is resistant to EndoH treatment

Because C6f1 was expressed in Lec1 cells producing high mannose glycans, these glycans should be sensitive to EndoH, an endoglycosidase that cleaves within the chiobiose core of high-mannose structures. To test whether treatment with this enzyme would remove the MAb109 epitope from C6f1 expressed in HEK-Lec1 cells, the glycoprotein was treated EndoH and tested for MAb109 reactivity following SDS-PAGE (Fig. 2.17). Surprisingly, the C6f1 remained MAb109-reactive after EndoH treatment, although other glycans on the Lec1-derived C6f1 were removed by the glycosidase. A shift in molecular weight of approximately 13 kDa was observed after EndoH treatment and this shift in molecular weight was observed for both MAb109 and anti-Fc reactivity (left and center panel of Fig. 2.17). The E-PHA lectin blot shown in the right panel was absent for E-PHA reactivity, which is consistent with previous findings that the MAb109 product was not a β1,4-linked GlcNAc. Taken together, these results demonstrate that the MAb109 epitope modification causes the epitope to be resistant to EndoH activity.

Further characterization of C6f1 glycans expressed in HEK-Lec1 cells was performed

using mass spectrometry. A mass spectra of permethylated glycans after PNGaseF treatment of C6f1 showed the presence of the high-mannose series of glycans with m/z peaks being observed for Man_5-, Man_6-, Man_7-, Man_8-, and Man_9-$GlcNAc_2$ glycans (Fig. 2.18, upper panel). Man_5- was the most prevalent glycan when compared against the internal standards, 3-O-β-cellobiosyl-D-glucose (DP3) and 3-O-β-cellotriosyl-D-glucose (DP4), which were spiked into the sample, and was present at two different m/z (1+ and 2+). The signal intensity for Man_5Glc_2 (1+) was approximately ten-fold the signal intensity of the standards. These results were consistent with glycans synthesized by cells lacking GnT-I activity which cannot synthesize hybrid or complex glycans. If the MAb109 was an additional HexNAc, as previously predicted by MS analysis of MAb109-reactive glycans in HEK-293 cells, a peak corresponding to $Man_5HexNAc_3$ should have been observed; however, this peak was not observed because MAb109 eptiope is added to a minority of all glycans expressed, and the prevalent high-mannose glycans obscured the signal corresponding to $Man_5HexNAc_3$. C6f1 was treated with EndoH before PNGaseF was used to release glycans in attempts to remove the prevalent, non-MAb109-reactive high-mannose glycans before detecting the MAb109-reactive glycan.

MS spectra of C6f1 glycans treated with EndoH before PNGaseF release, permethylation, and analysis showed that the peaks corresponding to Man_5-, Man_6-, Man_7-, Man_8-, and Man_9-$GlcNAc_2$ glycans were absent after EndoH treatment (Fig. 2.18, lower panel). This observation was consistent with the shift in molecular weight observed for MAb109 immunoblots of EndoH treated C6f following SDS-PAGE. A minor peak at m/z 907 had the approximate mass of $Man_5HexNAc_3$ and further supported the hypothesis that the MAb109 epitope containing an additional HexNAc

(Fig. 2.18, lower panel). The signal of this peak was only 20% of the signal of the standard and was not enough material to perform MS^n analysis; however, that presence of a terminal HexNAc was likely responsible for MAb109 reactivity was supported by a TIM trace.

NMR (Nuclear magnetic resonance) analysis of MAb109-reactive glycan isolated from HEK-Lec1 cells. Previous structural studies analyzing MAb109-reactive glycans isolated from HEK-293 cells by mass spectrometry had identified an additional HexNAc bound to the β-linked mannose of an N-linked glycan. Because neither the identity, nor the linkage of this sugar, could be identified by MS, the MAb109 needed to be analyzed by NMR, a technique that could identify the sugar, linkage, and α or β conformation. MAb109-reactive glycans were isolated from C6f1 expressed in HEK-Lec1 cells (Fig. 2.19) by removing non-MAb109-reactive glycans before releasing MAb109-reactive glycans with PNGaseF. Both one-dimensional and two-dimensional spectra for the MAb109-reactive glycans were obtained using a 600 MHz instrument. In both spectra, many peaks could be observed between 4.0 and 5.5 ppm, an area where carbohydrates are commonly detected (Fig. 2.20). The α-linked mannoses were detected at 4.9 and 5.1 ppm, while the β-linked mannose was observed at 4.8. Additionally, the H2 cross peaks were observed on the two-dimensional spectra. Although, the mannoses of the glycan expressing the MAb109 were detected, the HexNAc could not be identified. A peak at approximately 4.45 shows characteristics of both GalNAc and GlcNAc was observed, but could not be definitely assigned. Additional MAb109-reactive glycans

have been isolated and await analysis.

MAb109 epitope addition requires a specific N-linked glycan site and C-terminal amino acid sequence of CEACAM6. The observation that the MAb109 epitope is added to a specific N-linked site of a small set of glycoproteins, perhaps as small as two, led us to hypothesize that addition of the epitope required protein-specific sequence determinants. To test this hypothesis and identify these determinant regions, smaller cDNA encoding fragments common to CEACAM6 and C6f1 were PCR- amplified, ligated into the pFUSE vector, and expressed in HEK-293 cells (Fig. 2.21). A schematic of the various CEACAM6 fragment (C6f1-C6f6) cDNAs in spatial relation to the full-length CEACAM6 is shown in Figure 2.21A. CEACAM6 consists of three immunoglobulin domains and a GPI-anchor shown in relation to the full-length CEACAM6 cDNA. The numbers below the Ig-domain represent the corresponding amino acid numbers of CEACAM6 as well as the total number of amino acids for that domain in parenthesis. The CEACAM6 cDNA encoding fragments, labeled C6f2 through C6f6, were constructed, taking into consideration the N-link site identified as occupied by the MAb109 epitope and the Ig-domain constraints.

After transient transfection using HEK-293 cells, only one cDNA, C6f5, which contained both the N-linked site that expresses the epitope and downstream amino acid sequence, demonstrated MAb109 reactivity (Fig. 2.21B). C6f5 was secreted as a 58 kDa band when HEK supernatants were subjected to SDS-PAGE and immunoblotted with MAb109. Expression of the epitope was not observed, however, with any of the other clones, which contained only the N-linked site that expresses the epitope or this

site in combination with N-terminal peptide sequence. The lack of MAb109 reactivity of the fragments was due to lack of epitope expression and not from lack of glycoptrotein expression since anti-Fc reactivity was positive for all five clones (Fig. 2.21B). These results suggest that both the N-linked site that expresses the epitope and sequences downstream of the occupied N-linked site are required for the expression of the epitope.

MAb109 expressing and non-expressing proteins have C-terminal cluster of non-homologous amino acids. To identify the specific regions of the C-terminus that convey the protein-specific addition of the MAb109 epitope, we performed amino acid sequence alignment of two proteins known to express the MAb109 epitope, CEACAM5 and CEACAM6, along with a closely related CEACAM member that does not express the epitope, CEACAM8 (Fig. 2.22). The alignment of amino acids 191 to 318, which correspond to the amino acids of the smallest cDNA fragment that expressed the MAb109 epitope, revealed that the peptide sequences adjacent to the N-linked site that expressed the epitope were nearly identical between MAb109-reactive and non-reactive proteins. Interestingly, upon searching for sequence differences in the C-terminus of the sequences, we identified three consecutive amino acids that were clearly different between the CEACAM5 and CEACAM6 sequences and the CEACAM8 sequence. We hypothesized that these amino acids, 300-302, were necessary for determining addition of the MAb109 epitope to specific glycoproteins.

Specific amino acids in the C-terminus of CEACAM6 are required for MAb109 epitope addition. The sequence alignment of MAb109-reactive and non-reactive proteins

identified a cluster of amino acids that were clearly different between the two groups. To test whether the amino acids identified by sequence alignment were required for addition of the epitope, the CEACAM6, ^{300}QAH^{302}, sequence was mutated to that of the non-MAb109 epitope-expressing CEACAM8 sequence, ^{300}HTT^{302}. The mutated C6f1 lacked significant MAb109 epitope expression when secreted from HEK cells (Fig. 2.23, left panel). The large spread in molecular weight of the MAb109-reactive C6f1 from approximate 50-95 kDa was due to overexposure of the X-ray film in attempts to detect MAb109 reactivity of the mutated C6f1. The MAb109 epitope was not added to the mutated C6f1 construct despite its expression at levels nearly equal to non-mutated C6f1, as measured by anti-Fc reactivity (Fig. 2.23, right panel). The ^{300}QAH^{302} amino acids of CEACAM6 appear to act as sequence recognition determinants for the glycosyltransferase or glycosyltransferases that adds the MAb109 epitope. This conclusion would be further confirmed if addition of the specific sequences to a non-MAb109 expressing protein resulted in addition of the epitope.

Therefore, the CEACAM8 sequence, ^{300}HTT^{302}, was mutated to that of the MAb109 epitope-expressing CEACAM6 sequence, ^{300}QAH^{302}. Mutation of these three amino acids was sufficient to express a CEACAM8 glycopeptide positive for MAb109 reactivity in HEK cells, whereas the non-mutated CEACAM8 fragment did not (Fig. 2.24, left panel). Both glycoprotein fragments were expressed at nearly equal levels (Fig. 2.24, right panel). These data further support the hypothesis that the glycosyltransferase or glycosyltransferases responsible for MAb109 epitope addition recognize a sequence determinant downstream from the N-linked site that expresses the epitope. These

results likely explain the observation that MAb109 epitope addition occurs only a small subset of glycoproteins.

DISCUSSION

Much of the beneficial impact of a pancreatic cancer biomarker arises from its ability to detect pre-malignant or early stages of pancreatic adenocarcinoma (PDAC) when clinical intervention is still possible. Current biomarkers, which detect PDAC in late stages or after metastasis has occurred, are not useful clinically; therefore, the identification of new biomarkers for pancreatic cancer is of supreme importance. These studies have identified a novel, N-linked glycan expressed on CEACAM6 in pancreatic cancer reacts with MAb109, and this epitope is a potential biomarker for the diagnosis and treatment of pancreatic adenocarcinoma. Furthermore, addition of the MAb109 epitope is limited to specific N-linked sites of glycoproteins expressing sequence determinants.

MAb109 expressing glycoproteins, CEACAM5 and CEACAM6, have been well characterized in the literature. CEACAM5, also known as carcinoembryonic antigen (CEA), was first described nearly fifty years ago as a 180 kDA glycoprotein overexpressed in the fetal colon and colonic carcinomas [19]. An increased serum CEACAM5 level proved to be a clinically useful marker for colon adenocarcinoma and has been adopted to indicate colon cancer tumor burden and measure treatment effectiveness [20]. Molecular cloning of CEACAM5 revealed that it belonged to a gene family consisting of 29 genes found on chromosome 19. Other members of this family include pregnancy specific glycoproteins (PSGs), biliary glycoprotein (BGP), and

additional CEACAM family members [21-23]. A member of the CEACAM family, CEACAM6, was originally identified as having serological cross-reactivity to CEACAM5 [24]. Both CEACAM5 and CEACAM6 express a glycophosphatidylinositol (GPI) anchor, immunoglobulin domains, and N-linked glycans as well as perform common functions including cell adhesion and inhibition of cell differentiation. Despite the commonalities of both glycoproteins, CEACAM5 and CEACAM6 have different transcription factor binding sites and regulatory elements within their promoters [25, 26]. These variances may explain why the MAb109 epitope is expressed on CEACAM6 and is a marker for pancreatic cancer.

Our data establish that the MAb109 epitope expressed in pancreatic cancer is an N-linked glycan that is expressed on CEACAM6. Treating a pancreatic cancer cell line, BxPC3, with PNGaseF, we observed that the epitope was sensitive and that MAb109 reactivity was completely abolished compared to an untreated control (Fig. 2.1). Competitive inhibition assays using C6f1 glycoprotein, glycopeptides, and glycans further demonstrated that the MAb109 recognizes a glycan specifically expressed on CEACAM6. MAb109 bound to C6f1 glycoprotein, C6f1 glycopeptides, and C6f1 glycans, but not to another known glycoprotein RNaseB, resulting in inhibition on C6f1 dot plots. Additionally, the C6f1 glycopeptides had been isolated after C6f1 treatment with trypsin and chymotrypsin providing evidence that MAb109 recognition of its epitope does not require C6f1 secondary or tertiary structure. Even though PNGaseF abolished MAb109 reactivity, it is not possible to completely rule out an affinity of MAb109 towards C6f1 peptide (Fig. 2.9). Additional experiments such as surface plasmon resonance (SPR)

would allow for the calculation of binding affinities of the MAb109 toward glycan and CEACAM6 peptide.

Although the MAb109 epitope is endogenously expressed on CEACAM6 in pancreatic cancer, the epitope can also be expressed *in vitro* by additional cell lines. In this study, every mammalian cell line that we transfected with the C6f1 construct expressed the MAb109 epitope. Many of these cells were negative for both CEACAM6 and MAb109, however, the cells were positive for MAb109 after transfection with the C6f1 construct. These findings indicate that these cells contained the biosynthetic machinery to express the MAb109 despite not expressing the required acceptor. Presumably, the enzyme or enzymes required for MAb109 epitope addition act when presented with the correct acceptor. The observance that cells expend energy to express the biosynthetic machinery to add the MAb109 epitope to an acceptor, which may or may not be present, suggests that the epitope may be evolutionarily conserved and have a critical function.

Currently, the functional significance of the MAb109 epitope is not understood. Previous studies have provided insight into the function of the CEACAM6, including maintenance of cell polarity, tissue architecture, and cell differentiation [27, 28]. Overexpression of CEACAM6 during tumorigenesis has been shown to decrease cell adhesion, increase cell invasiveness, and cause resistance to anoikis [28-31]. Therefore, it is not difficult to invision how tumor cells would gain survival advantages by overexpressing CEACAM6. Although previous studies investigating CEACAM6 during tumorigenesis have been performed, the role of N-linked glycans, especially as it pertains MAb109 expression, needs to be investigated. Since the epitope is

overexpressed during development and oncogenesis, conditions favoring cell plasticity, the MAb109 epitope could have a role in decreasing cell adhesion properties of CEACAM6. This hypothesis was supported by the observation that the MAb109 epitope was found specifically on Asp^{197} of CEACAM which is found in a region known to mediate cell adhesion effects [32].

Investigating the function of the MAb109 epitope will require the identification of the glycosyltransferase or glycosyltransferases responsible for its addition. Only three mammalian glycosyltransferases have been previously described to add sugar residues to the trisaccharide core of an N-linked glycan. FucT-VIII transfers an α1,6-linked fucose to the proximal GlcNAc residue adjacent to asparagine [33]. Also, two additional glycosyltransfereases, GlcNAcT-VII and VIII, have been described by Stanley's laboratory in mutagenized CHO cells [34-36]. GlcNAcT-VII catalyzes the addition of a β1,2 linked GlcNAc to the β-linked mannose residue [34, 36], and GlcNAcT-VIII adds a β1,6 linked GlcNAc to the core GlcNAc residue distal to asparagine attachment [34, 35]. It is unlikely, however, that one of these three enzymes is involved in the expression of the epitope. MS and NMR data showed that the epitope is a HexNAc linked to the β-linked mannose. This finding eliminates Fut-VIII, which transfers fucose, and GnT-VIII, which transfers GlcNAc to the chitobiose core. GnT-VII does add a HexNAc to the β-linked mannose, but Stanley's lab showed that the preferred acceptor for this enzyme is a biantennary glycan acceptor terminating in galactose with core fucosylation. We did not observe fucose on any of the glycans expressing the MAb109 epitope; however, GnT-VII cannot be ruled out since it was first identified in a mutated cell line. A mutation may have been introduced which caused the enzyme to preferentially recognize a

biantennary glycan with fucose. Therefore, the synthesis of the MAb109 requires a currently unidentified glycosyltransferase with an activity not yet described or GnT-VII from mutagenized CHO cells shows alter acceptor specificity. Determining the exact structure of the N-linked glycoepitope should suggest possible enzyme activities required for MAb109 epitope biosynthesis.

The glycosyltransferase or glycosyltransferases identified to add the MAb109 epitope will likely contain regions that recognize and modify specific proteins. Specific amino acid sequence determinants found in CEACAM6 that were required for addition of the MAb109 epitope. This observation was consistent with other examples in the literature in which a specific glycan or glycans are expressed only at particular sites of specific proteins as the glycosyltransferase requires sequence determinants in the underlying glycoprotein acceptor. For example, the addition of β1,4-GalNAc to terminal GalNAc by β1,4–N–acetylgalactosaminyltransferase 3 and 4 (βGT3/4) occurs only on glycoprotein hormones such as luetenizing hormone (LH) and human chorionic gonadotropin (hCG). These proteins contain a Pro-X-Lys motif with the lysine residues appropriately spaced and positioned relative to the N-glycan being modified have been shown to be important determinants of acceptor specificity [37]. Additional examples of protein specific glycosylation include the specific N-acetylglucosamine (GlcNAc)-1-phosphate addition on terminal mannose residues of lysosomal enzymes, which requires lysine residues appropriately spaced and positioned relative to the N-glycan being modified, and the O-fucosylation of correctly folded cysteine-rich domains found in epidermal growth factor (EGF) and thrombospondin repeats (TSR) [38-40].

The polysialylation of the neural cell adhesion molecule (NCAM) is the most well studied example of protein-specific glycosylation and shares the most homology to the MAb109 epitope addition to CEACAM6 (Fig. 2.25). Polysialylation is a protein-specific posttranslational modification that results in the addition of α2,8-linked sialic acid polymer by two glycosyltransferase, PST or STX, to an N-linked glycan site found on the fifth Ig-domain of NCAM. Polysialic acid addition is guided by glycosyltransferases recognizing a unique amino acid sequence between the strands 4 and 5 of its ß sandwich structure in the type III fibronectin domain adjacent to the fifth Ig-domain [41]. This observation is consistent with our data where we identified a three amino acid group common to MAb109-reactive glycoproteins that was required for epitope expression on Asp^{197} in an adjacent Ig-domain. This sequence was clearly different in non-reactive glycoproteins. The amino acid sequence likely act as a recognition domain for the currently unidentified glycosyltransferase to specifically bind to CECAM6, but not to additional glycoproteins, and add the MAb109 eptiope. Additionally, removal or mutation of the specific, C-terminal amino acid sequences results in the absence the unique, N-linked glycan expressed independently on NCAM and CEACAM6.

Based on the finding that the addition of the MAb109 epitope requires specific sequence determinants in relation to the occupied N-linked site, it is plausible that the epitope may be expressed on additional glycoproteins besides CEACAM5 and CEACAM6. Thus, we would predict that MAb109 epitope would be expressed on a glycoprotein containing both Ig-domains and at least one predicted N-linked sequon. This prediction would be consistent with our findings for the sequence determinants required for addition of the MAb109 epitope. Recently, preliminary studies with pupae

from Drosophila have identified a MAb109-reactive band on Western blot at ~75 kDa after SDS-PAGE. This is likely a previously unidentified acceptor since insects do not express CEACAM5 or CEACAM6; however, the acceptor is likely another Ig-family glycoprotein, which are known to be expressed in Drosophila [42]. Additionally, MAb109 epitope exhibits stage specific expression in the pupae when the digestive tract of the fly is undergoing development, which is consistent with MAb109 being expressed on an oncofetal protein having a role in patterning and cell architecture. Identifying additional MAb109 epitope expressing proteins would potentially provide insight into the function of the epitope as well as its prevalence.

Not much is known about the prevalence of the MAb109 epitope or about additional glycans or glycoproteins that could potentially be used as cancer biomarkers. Glycan biomarkers are not abundant making them difficult to identify as a novel structure instead of one previously identified. For example, in this study, the MAb109 epitope would have undoubtedly been misidentified as the common β1,2-linked GlcNAc in MS based glycoproteomic studies without the specificity of the antibody. As is the case with this example and many others, the discovery of specific glyomarkers requires the use of monoclonal antibodies [43]. Potential for identifying these the techniques for isolating minor glycan population, identification of additional glycan structures for biomarkers as the technologies required for analyzing structure, cataloging these structures improves.

Even though MAb109 epitope expression on CEACAM6 is first example of protein-specific glycan being used as a cancer biomarker, other examples are currently under investigation. Another monoclonal antibody produced and isolated by Cohava Gelber, VAC69, shows specificity for multiple myeloma (MM) and ovarian carcinoma, but does

not bind to non-cancerous cells and tissues tested. VAC69 was produced against MM cells after tolerization to a related cell type, and in five MM cell lines, the VAC69 epitope was expressed on a single glycoprotein with approximate molecular weight of 80 kDa. Preliminary results showed that PNGaseF (N-glycanase) treatment of MM cell lysates resulted in the loss of VAC69 binding to Western blots of MM cell lysates. This observation demonstrates that the epitope was most likely part of an N-glycan or N-glycopeptide. Despite both antibodies demonstrating reactivity towards N-linked glycans, MAb109 and VAC69 did not cross-react when assayed on cells or Western blots of lysates after SDS-PAGE (data not shown). The specificity of the antibody suggests that the VAC69 epitope, like the MAb109 epitope, may be a novel glycan added to a specific subset of glycoprotein containing sequence determinants. Furthermore, VAC69 provides another example of glycans being possible biomarkers for cancer and that antibodies against these glycans could be used to develop clinical diagnostics and therapeutics.

Our current findings raise the possibility of exploiting the MAb109 epitope as a biomarker for the diagnosis and treatment of pancreatic adenocarcinoma. IHC staining of pancreatic tissue sections successfully identified dysplastic pancreas, an early stage of cancer development in which curative therapy could occur. Additionally, preliminary studies have shown that MAb109 epitope can be detected on Western blot of pancreatic ductal fluids (PDF) taken from pancreatic adenocarcinoma patients but not non-diseased normals (data not shown). Additional PDF samples with controls need to be analyzed to increase the sample number in order to determine sensitivity and specificity of the antibody, but these studies set the stage for the potential development of a

clinical diagnostic test. Since CEACAM6 is a GPI-anchored protein which could be cleaved by phosphilipases, we anticipate that CEACAM6 expressing the MAb109 epitope will be in the bloodstream where it could be detected. This would allow the development of a non-invasive serum blood test for pancreatic cancer, a test currently sorely needed in the clinic. In addition to a diagnostic test, MAb109 could potentially be linked to a pharmacological agent in order to develop a monoclonal antibody based therapeutic.

In conclusion, our data indicates that the MAb109 epitope is a novel, glycan that is expressed on a specific, N-linked site on CEACAM6 in PDAC. This glycan is protein specific, and its addition requires sequence determinants found in the C-terminus of CEACAM6. Therefore, the identification of the MAb109 epitope provides another example of protein-specific glycosylation of a novel glycan, but also can foster the development of novel diagnostics and therapeutics for PDAC.

REFERENCES

1. *American Cancer Society*. Cancer Facts & Figures, 2014.
2. Oberstein, P.E. and M.W. Saif, *First-line treatment for advanced pancreatic cancer. Highlights from the "2011 ASCO Gastrointestinal Cancers Symposium". San Francisco, CA, USA. January 20-22, 2011.* JOP, 2011. **12**(2): p. 96-100.
3. Zhu, C.P., et al., *Gemcitabine in the chemoradiotherapy for locally advanced pancreatic cancer: a meta-analysis.* Radiother Oncol, 2011. **99**(2): p. 108-13.
4. Howlader, N., et al., *Improved estimates of cancer-specific survival rates from population-based data.* J Natl Cancer Inst, 2010. **102**(20): p. 1584-98.
5. Duxbury, M.S., et al., *CEACAM6 is a novel biomarker in pancreatic adenocarcinoma and PanIN lesions.* Ann Surg, 2005. **241**(3): p. 491-6.
6. Spiro, R.G., *Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds.* Glycobiology, 2002. **12**(4): p. 43R-56R.
7. Hwa, K.Y., *Glycosyl phosphatidylinositol-linked glycoconjugates: structure, biosynthesis and function.* Adv Exp Med Biol, 2001. **491**: p. 207-14.
8. Hildebrandt, H., et al., *Dissecting polysialic acid and NCAM functions in brain development.* J Neurochem, 2007. **103 Suppl 1**: p. 56-64.
9. Wells, L., S.A. Whelan, and G.W. Hart, *O-GlcNAc: a regulatory post-translational modification.* Biochem Biophys Res Commun, 2003. **302**(3): p. 435-41.
10. Wasley, L.C., et al., *The importance of N- and O-linked oligosaccharides for the biosynthesis and in vitro and in vivo biologic activities of erythropoietin.* Blood, 1991. **77**(12): p. 2624-32.
11. Ohtsubo, K. and J.D. Marth, *Glycosylation in cellular mechanisms of health and disease.* Cell, 2006. **126**(5): p. 855-67.
12. Buckhaults, P., et al., *Transcriptional regulation of N-acetylglucosaminyltransferase V by the src oncogene.* J Biol Chem, 1997. **272**(31): p. 19575-81.
13. Kang, R., et al., *Transcriptional regulation of the N-acetylglucosaminyltransferase V gene in human bile duct carcinoma cells (HuCC-T1) is mediated by Ets-1.* J Biol Chem, 1996. **271**(43): p. 26706-12.
14. Alvarez-Manilla, G., et al., *Tools for glycomics: relative quantitation of glycans by isotopic permethylation using 13CH3I.* Glycobiology, 2007. **17**(7): p. 677-87.
15. Kannagi, R., *Carbohydrate antigen sialyl Lewis a--its pathophysiological significance and induction mechanism in cancer progression.* Chang Gung Med J, 2007. **30**(3): p. 189-209.
16. Chen, R., et al., *Glycophosphatidylinositol-anchored protein deficiency as a marker of mutator phenotypes in cancer.* Cancer Res, 2001. **61**(2): p. 654-8.
17. Malykh, Y.N., R. Schauer, and L. Shaw, *N-Glycolylneuraminic acid in human tumours.* Biochimie, 2001. **83**(7): p. 623-34.
18. Pierce, J.M., *Cancer Glycomics. Handbook of Glycomics*, ed. Cummings. 2009, San Diego: Academic Press.
19. Gold, P. and S.O. Freedman, *Demonstration of Tumor-Specific Antigens in Human Colonic Carcinomata by Immunological Tolerance and Absorption Techniques.* J Exp Med, 1965. **121**: p. 439-62.

20. Moertel, C.G., et al., *An evaluation of the carcinoembryonic antigen (CEA) test for monitoring patients with resected colon cancer.* JAMA, 1993. **270**(8): p. 943-7.
21. Khan, W.N., et al., *The pregnancy-specific glycoprotein family of the immunoglobulin superfamily: identification of new members and estimation of family size.* Genomics, 1992. **12**(4): p. 780-7.
22. Olsen, A., et al., *Gene organization of the pregnancy-specific glycoprotein region on human chromosome 19: assembly and analysis of a 700-kb cosmid contig spanning the region.* Genomics, 1994. **23**(3): p. 659-68.
23. Thompson, J., et al., *Long-range chromosomal mapping of the carcinoembryonic antigen (CEA) gene family cluster.* Genomics, 1992. **12**(4): p. 761-72.
24. von Kleist, S., G. Chavanel, and P. Burtin, *Identification of an antigen from normal human tissue that crossreacts with the carcinoembryonic antigen.* Proc Natl Acad Sci U S A, 1972. **69**(9): p. 2492-4.
25. Koops, M.D., et al., *Transcriptional regulation of the non-specific cross-reacting antigen gene, a member of the carcinoembryonic antigen gene family up-regulated in colorectal carcinomas.* Eur J Biochem, 1998. **253**(3): p. 778-86.
26. Richards, C.A., E.A. Austin, and B.E. Huber, *Transcriptional regulatory sequences of carcinoembryonic antigen: identification and use with cytosine deaminase for tumor-specific gene therapy.* Hum Gene Ther, 1995. **6**(7): p. 881-93.
27. Poola, I., et al., *Expression of carcinoembryonic antigen cell adhesion molecule 6 oncoprotein in atypical ductal hyperplastic tissues is associated with the development of invasive breast cancer.* Clin Cancer Res, 2006. **12**(15): p. 4773-83.
28. Blumenthal, R.D., H.J. Hansen, and D.M. Goldenberg, *Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen).* Cancer Res, 2005. **65**(19): p. 8809-17.
29. Ordonez, C., et al., *Human carcinoembryonic antigen functions as a general inhibitor of anoikis.* Cancer Res, 2000. **60**(13): p. 3419-24.
30. Duxbury, M.S., et al., *CEACAM6 is a determinant of pancreatic adenocarcinoma cellular invasiveness.* Br J Cancer, 2004. **91**(7): p. 1384-90.
31. Duxbury, M.S., et al., *CEACAM6 gene silencing impairs anoikis resistance and in vivo metastatic ability of pancreatic adenocarcinoma cells.* Oncogene, 2004. **23**(2): p. 465-73.
32. Pavlopoulou, A. and A. Scorilas, *"A comprehensive phylogenetic and structural analysis of the carcinoembryonic antigen (CEA) gene family".* Genome Biol Evol, 2014.
33. Uozumi, N., et al., *Purification and cDNA cloning of porcine brain GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1-->6fucosyltransferase.* J Biol Chem, 1996. **271**(44): p. 27810-7.
34. Raju, T.S. and P. Stanley, *Gain-of-function Chinese hamster ovary mutants LEC18 and LEC14 each express a novel N-acetylglucosaminyltransferase activity.* J Biol Chem, 1998. **273**(23): p. 14090-8.
35. Raju, T.S., M.K. Ray, and P. Stanley, *LEC18, a dominant Chinese hamster ovary glycosylation mutant synthesizes N-linked carbohydrates with a novel core structure.* J Biol Chem, 1995. **270**(51): p. 30294-302.

36. Raju, T.S. and P. Stanley, *LEC14, a dominant Chinese hamster ovary glycosylation mutant expresses complex N-glycans with a new N-acetylglucosamine residue in the core region.* J Biol Chem, 1996. **271**(13): p. 7484-93.
37. Miller, E., et al., *A necessary and sufficient determinant for protein-selective glycosylation in vivo.* J Biol Chem, 2008. **283**(4): p. 1985-91.
38. Steet, R., W.S. Lee, and S. Kornfeld, *Identification of the minimal lysosomal enzyme recognition domain in cathepsin D.* J Biol Chem, 2005. **280**(39): p. 33318-23.
39. Ghosh, P., N.M. Dahms, and S. Kornfeld, *Mannose 6-phosphate receptors: new twists in the tale.* Nat Rev Mol Cell Biol, 2003. **4**(3): p. 202-12.
40. Leonhard-Melief, C. and R.S. Haltiwanger, *O-fucosylation of thrombospondin type 1 repeats.* Methods Enzymol, 2010. **480**: p. 401-16.
41. Mengeling, B.J., S.M. Manzella, and J.U. Baenziger, *A cluster of basic amino acids within an alpha-helix is essential for alpha-subunit recognition by the glycoprotein hormone N-acetylgalactosaminyltransferase.* Proc Natl Acad Sci U S A, 1995. **92**(2): p. 502-6.
42. Vogel, C., S.A. Teichmann, and C. Chothia, *The immunoglobulin superfamily in Drosophila melanogaster and Caenorhabditis elegans and the evolution of complexity.* Development, 2003. **130**(25): p. 6317-28.
43. Magnani, J.L., et al., *A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-N-fucopentaose II.* J Biol Chem, 1982. **257**(23): p. 14365-9.
44. Hammarstrom, S., *The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues.* Semin Cancer Biol, 1999. **9**(2): p. 67-81.

Figures

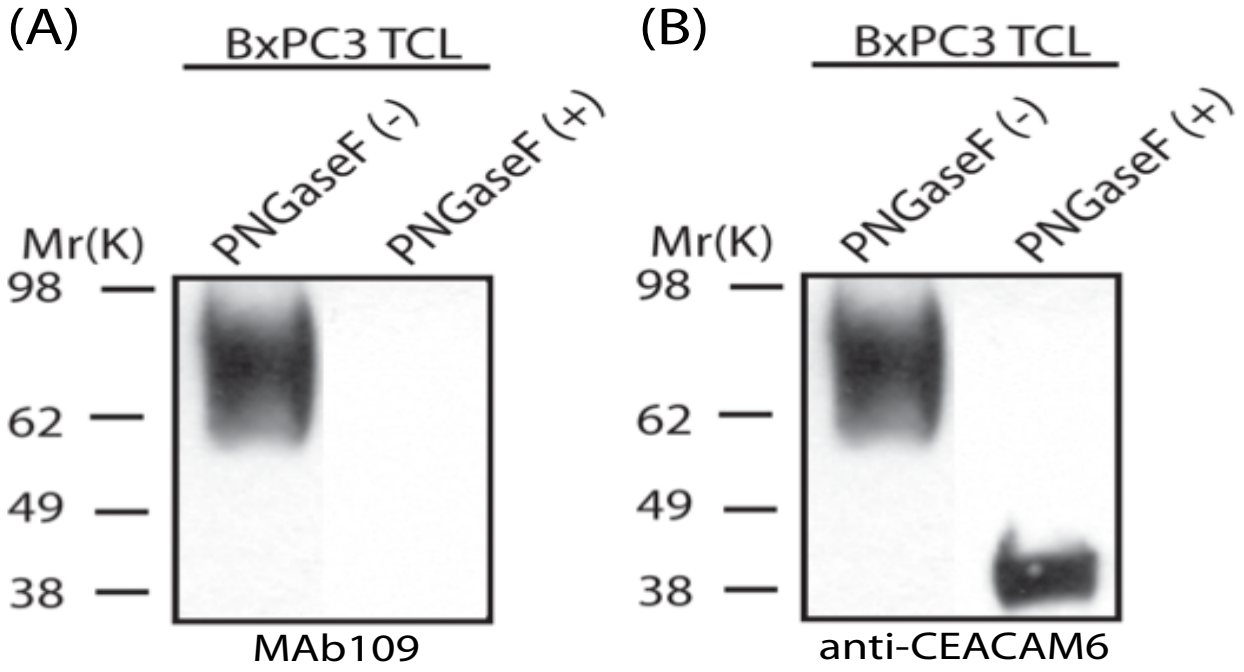


Figure 2.1. MAB109 REACTS WITH AN N-LINKED GLYCAN ON CEACAM6 EXPRESSED IN BXPC3 CELLS, A HUMAN PANCREATIC ADENOCARCINOMA (PDAC) CELL LINE. (A) SDS-PAGE of BxPC3 total cell lysates (TCL) followed by MAb109 immunoblotting revealed that the antibody reacted with a broad band with an approximate molecular weight centralized at 85 kDa. MAb109 reactivity was abolished after treatment of the BxPC3 total cell lysate with PNGaseF, a glycosidase which removes human N-linked glycans. (B) SDS-PAGE of BxPC3 total cell lysates followed by anti-CEACAM6 polypeptide immunoblotting shows that the protein was present despite the removal of N-linked glycans after PNGaseF treatment. PNGaseF treatment of BxPC3 total cell lysate caused an observed mass shift of approximately 45 kDa for the anti-CEACAM6 reactive band.

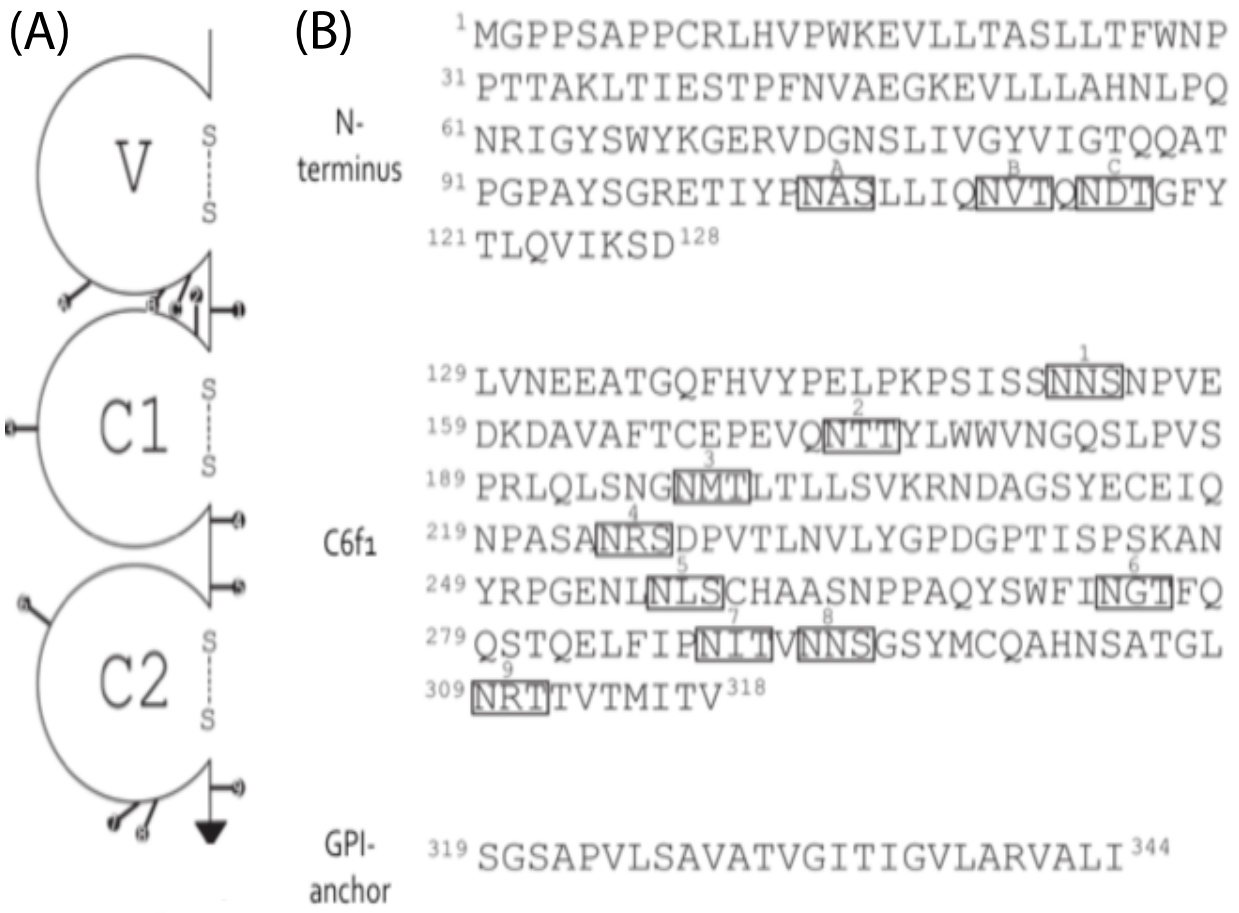


Figure 2.2. CEACAM6 IS A GLYCOPHOSPHATIDYLINOSITOL (GPI)-ANCHORED GLYCOPROTEIN THAT IS A MEMBER OF THE IMMUNOGLOBULIN (IG) SUPERFAMILY AND IS PREDICTED TO HAVE TWELVE N-LINKED SEQUONS. (A) Schematic of CEACAM6 molecule showing characteristic immunoglobulin domains and N-linked glycosylation sites (adapted from Hammarstrom [44]). CEACAM6 is a cell-surface glycoprotein that consists of three immunoglobulin domains, a N-terminal variable (V) domain and two C-terminal constant domains (C1 and C2), held together by disulfide bonds. CEACAM6 has 12 predicted N-linked sequons depicted as ball-and-stick extensions in panel A. (B) Amino acid sequence of CEACAM6 for the N-terminal variable domain (V), CEACAM6 fragment 1 (C6f1) consisting of the two C-type domains (C1 and C2), and the GPI-anchor. Predicted N-linked sites following the N-X-S/T motif are shown as enclosed boxes. In order to differentiate N-linked sites in the variable domain from those in C6f1, those sites found in the variable domain are lettered while those in C6f1 are numbered.

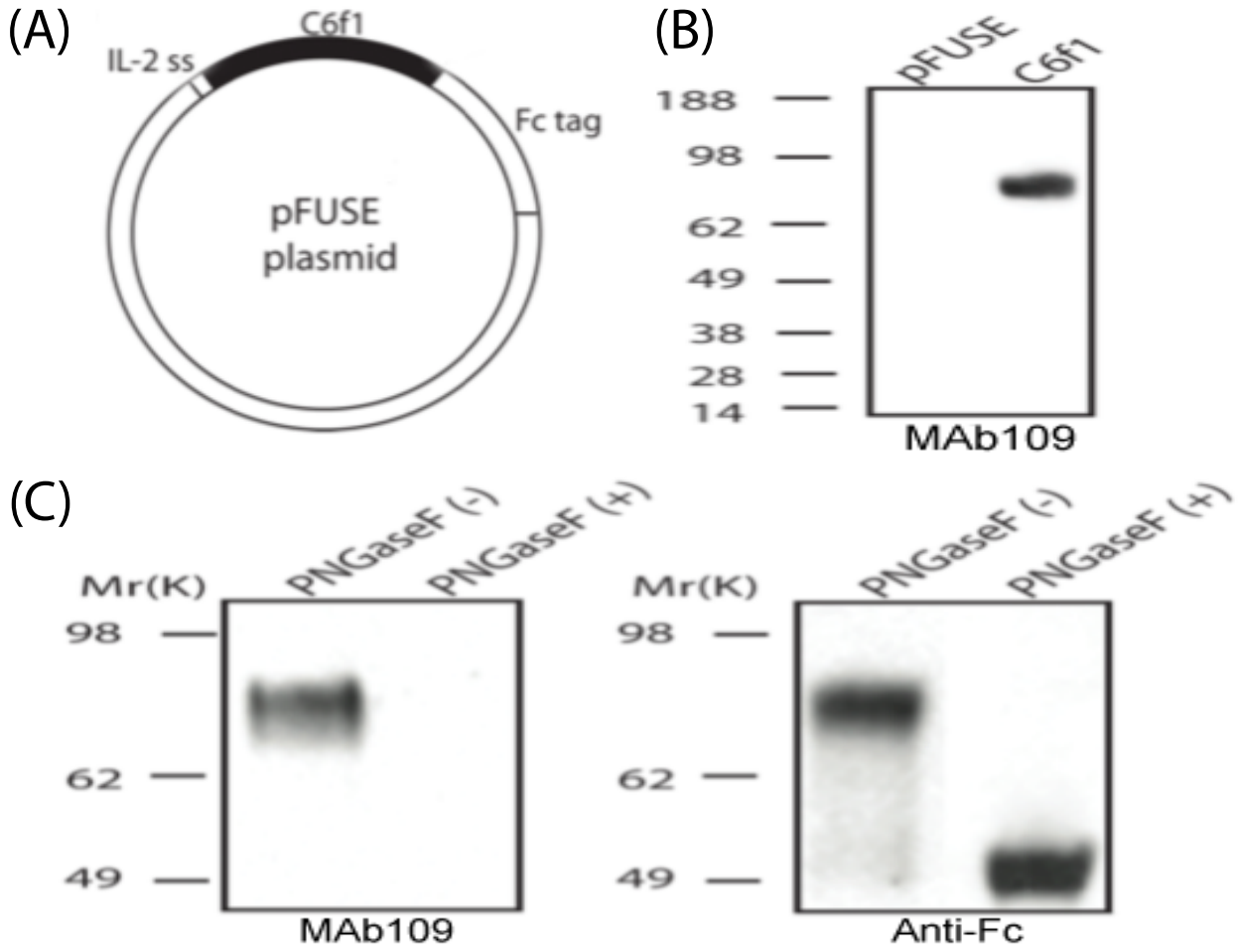


Figure 2.3. CEACAM6 FRAGMENT 1 (C6F1) TRANSIENTLY TRANSFECTED IN HEK-293 CELLS EXPRESSES THE MAB109 EPITOPE THAT IS PNGASEF SENSITIVE. (A) Graphical representation of C6f1 plasmid encoding, from N-terminus to C-terminus, IL-2 signal sequence, C6f1 insert, and Fc tag. (B) Immunoblots of supernatants taken from HEK-293 cells and subjected to SDS-PAGE show that C6f1 was expressed *in vitro* and was positive for MAb109 reactivity when transiently transfected with the C6f1 construct, but not the pFUSE control. (C) Left panel, C6f1 secreted from HEK-293 cells resulted in a single, 85 kDa band positive for MAb109 reactivity, but MAb109 reactivity was abolished after PNGaseF treatment. Right panel, anti-Fc antibody shows a 35 kDa shift corresponding to removal of N-linked glycans by PNGaseF. Anti-Fc reactivity in the PNGaseF (+) lane demonstrated the C6f1 polypeptide remained present despite removal of N-linked glycans and the lack of MAb109 reactivity.

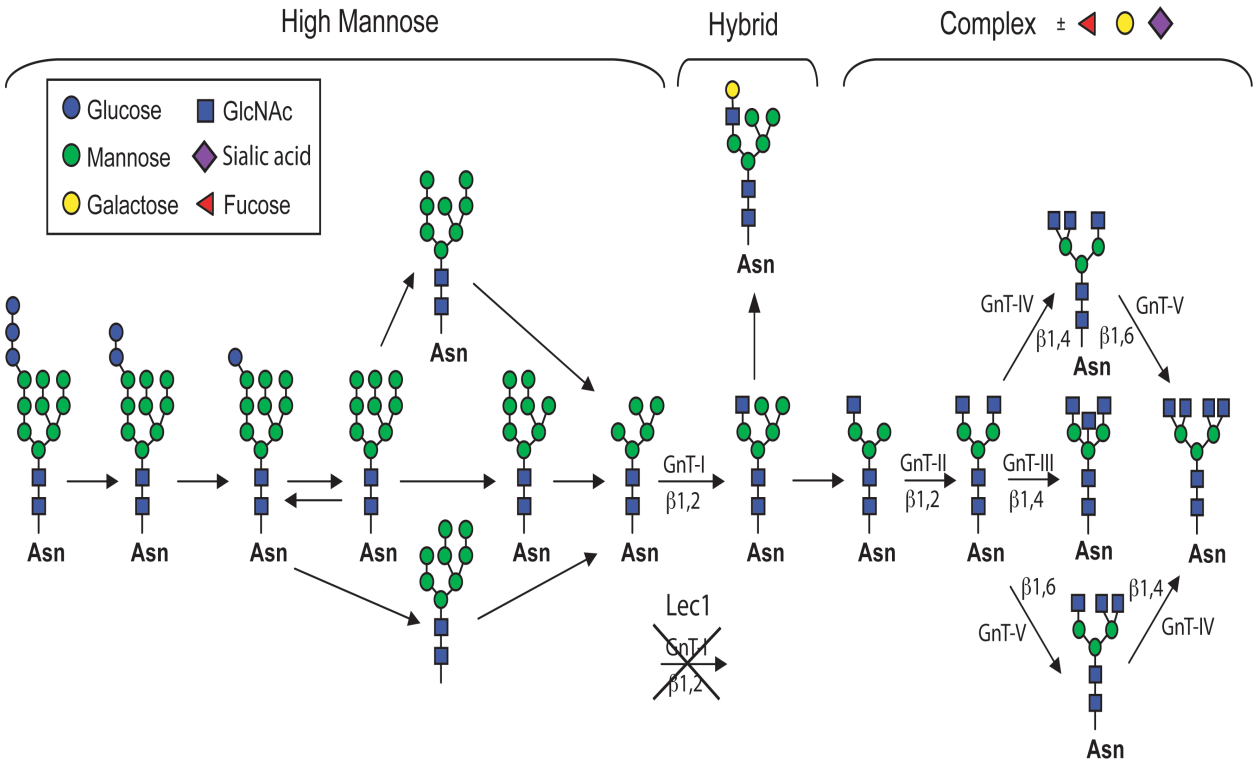


Figure 2.4. BIOSYNTHETIC PATHWAY OF N-LINKED GLYCAN BRANCHING. High-mannose, hybrid, and complex glycans are all synthesized from a common $Glc_3Man_9GlcNAc_2$ precursor that is added by oligosaccharyltransferase (OST) to newly synthesized proteins in the endoplasmic reticulum (ER). After addition of $Glc_3Man_9GlcNAc_2$, the glycan undergoes processing as terminal glucose residues are cleaved. Once correctly folded, the glycoprotein is trafficked through the ER, and the glycan undergoes successive mannose trimming by α-mannosidases, which ultimately yields a key intermediate, Man_5Glc_2. Prior to this intermediate, the glycans that contain only terminal mannose are known as high-mannose glycans. *N*-acetylglucosaminyltransferase-I (GnT-I) can add a GlcNAc residue to the C-2 of the α1-3 mannose in the core of Man_5Glc_2, which allows the biosynthesis of hybrid and complex glycans. It is important to note that hybrid and complex glycans cannot be synthesized without GnT-I activity. As a result, Lec1 cell lines that have GnT-I mutations only produce high-mannose glycans. This mutation is designated above with an X through the mutated GnT-I enzyme. Lec1 cells cannot synthesize N-linked glycans to the right of the mutation shown above.

Whether or not the subsequent GnT-I product, Man_5Glc_3 is acted on by α-mannosidase II will determine whether the glycan develops into a complex or hybrid glycan, respectively. If α-mannosidase II does act on the glycan, then GnT-II adds a β1-2 linked GlcNAc to the α1-6 linked mannose which creates a biantennary glycan, a structure that can be acted on by subsequent GlcNAc transferases to increase branching. The GlcNAc transferase and the linkage of the sugar that it adds are designated in the figure above. After addition of branching GlcNAc residues, each may be elongated with galactose, poly-*N*-acetyllactosamine, sialic acid, and fucose, except for the bisected GlcNAc, which is not extended.

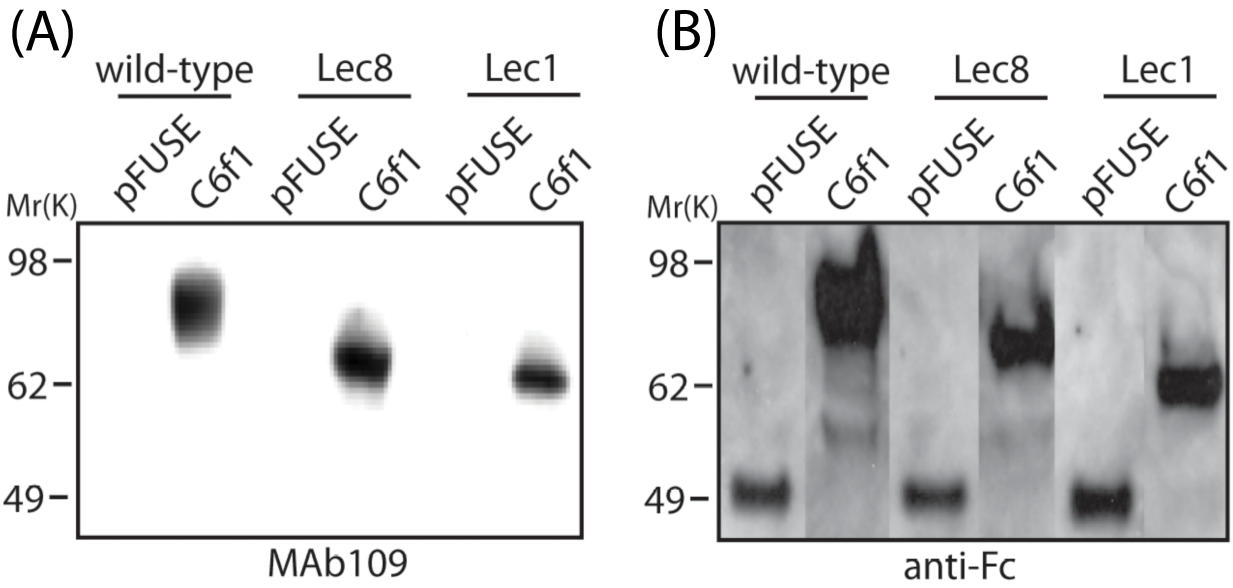


Figure 2.5. MAB109 REACTS WITH C6F1 RECOMBINANTLY EXPRESSED IN CHINESE HAMSTER OVARY (CHO) WILD-TYPE AND GLYCOSYLATION DEFICIENT CELL LINES. (A) SDS-PAGE and MAb109 immunoblots of supernatants collected from CHO wild-type, Lec8 (galactose transporter deficient), and Lec1 (GnT-I deficient) cell lines that had been transiently transfected with either the pFUSE or C6f1 construct. All three cell lines expressed MAb109 reactive C6f1 with the molecular weight of the MAb109 reactive band reflecting the glycosylation deficiency. (B) Immunoblot with anti-Fc antibody showing expression of recombinant constructs in all CHO cell lines.

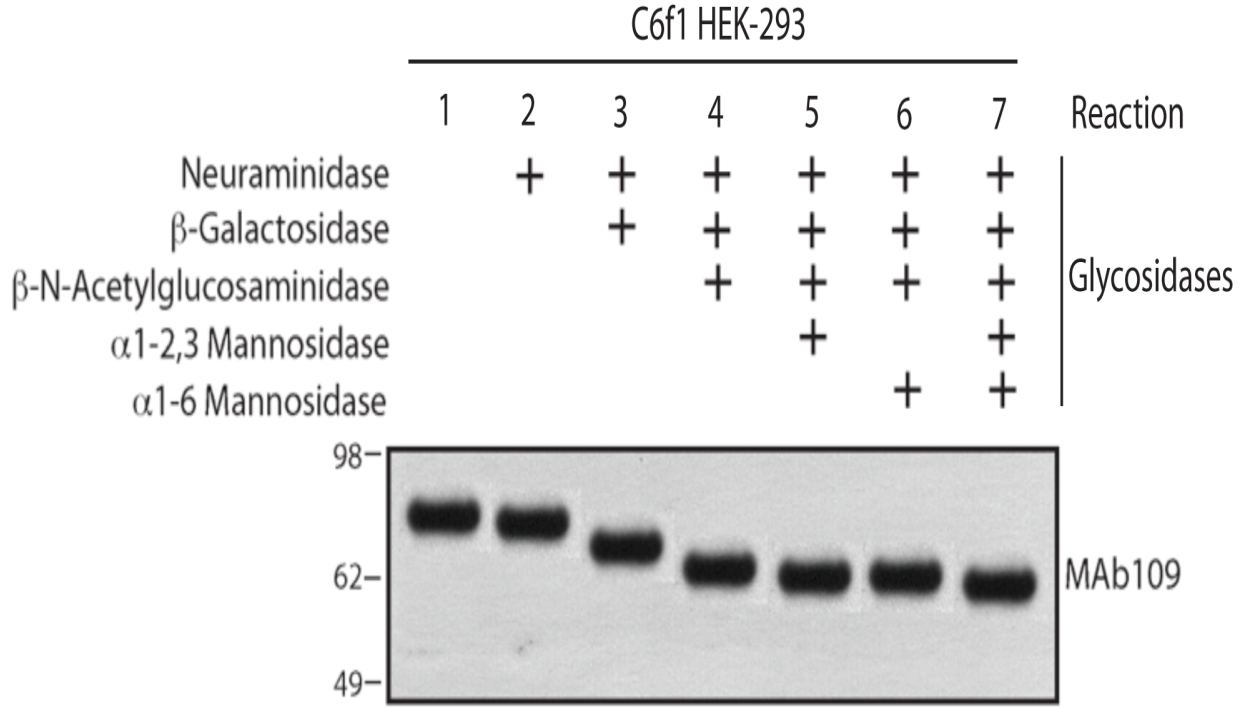


Figure 2.6. THE MAB109 EPITOPE IS RESISTANT TO GLYCOSIDASE TREATMENT. Recombinantly expressed C6f1 purified from HEK-293 cells was digested using combinations of neuraminidase, β-galactosidase, β-*N*-acetylglucosaminidase, α1-2,3 mannosidase, and α1-6 mannosidase, which are glycosidases that remove sialic acid, β-galactose, β-N-acetylglucosamine, and α-mannose from N-linked glycans, respectively. After digestion, the products of reactions 1 through 7 were subjected to SDS-PAGE and immunobloting with MAb109. Shifts to progressively lower molecular weight were observed with successive glycosidase treatments; however, MAb109 reactivity remained constant across all reactions despite the glycosidases removing terminal sugar residues from N-linked glycans.

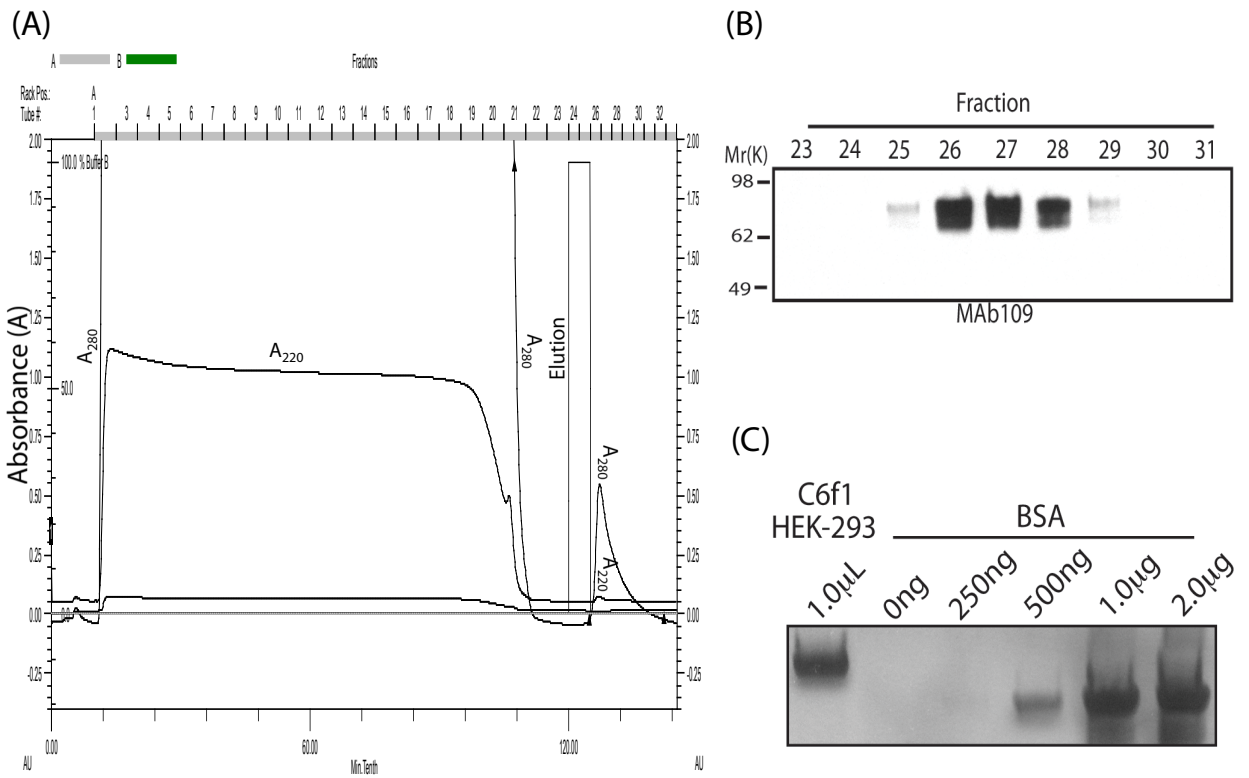


Figure 2.7. MILLIGRAM QUANTITIES OF RECOMBINANT C6F1 CAN BE PURIFIED USINGTRANSIENTLY TRANSFECTED HEK-293 CELLS. (A) Chromatograph of supernatant taken from HEK-293 cells transiently transfected with C6f1 and applied to immobilized protein A column. Column was washed with 20mM sodium phosphate buffer pH 7.0 (buffer A) before 150mLs of cell supernatant was bound at a flow rate of 5 mL/minute. Column was washed with 10 column volumes of buffer A before being eluted with 5 column volumes of 100mM citric acid pH 3.0 (buffer B). Elution of C6f1 was monitored by absorbance at 220 nm and 280 nm. Fractions predicted to contain purified C6f1 expressing the MAb109 eptiotpe were collected and immunoblotted with MAb109 following SDS-PAGE. (B) Immunoblot using MAb109 against fractions taken after low pH elution of protein A column. A 85kDa MAb109-reactive C6f1 band was present in the fractions corresponding to increased absorbance at 220nm (C) Silver stain of purified, C6f1 expressed in HEK-293 cells. Amount of recombinant C6f1 was calculated by comparing C6f1 staining against silver-stained bovine serum albumin (BSA) standards on the same gel.

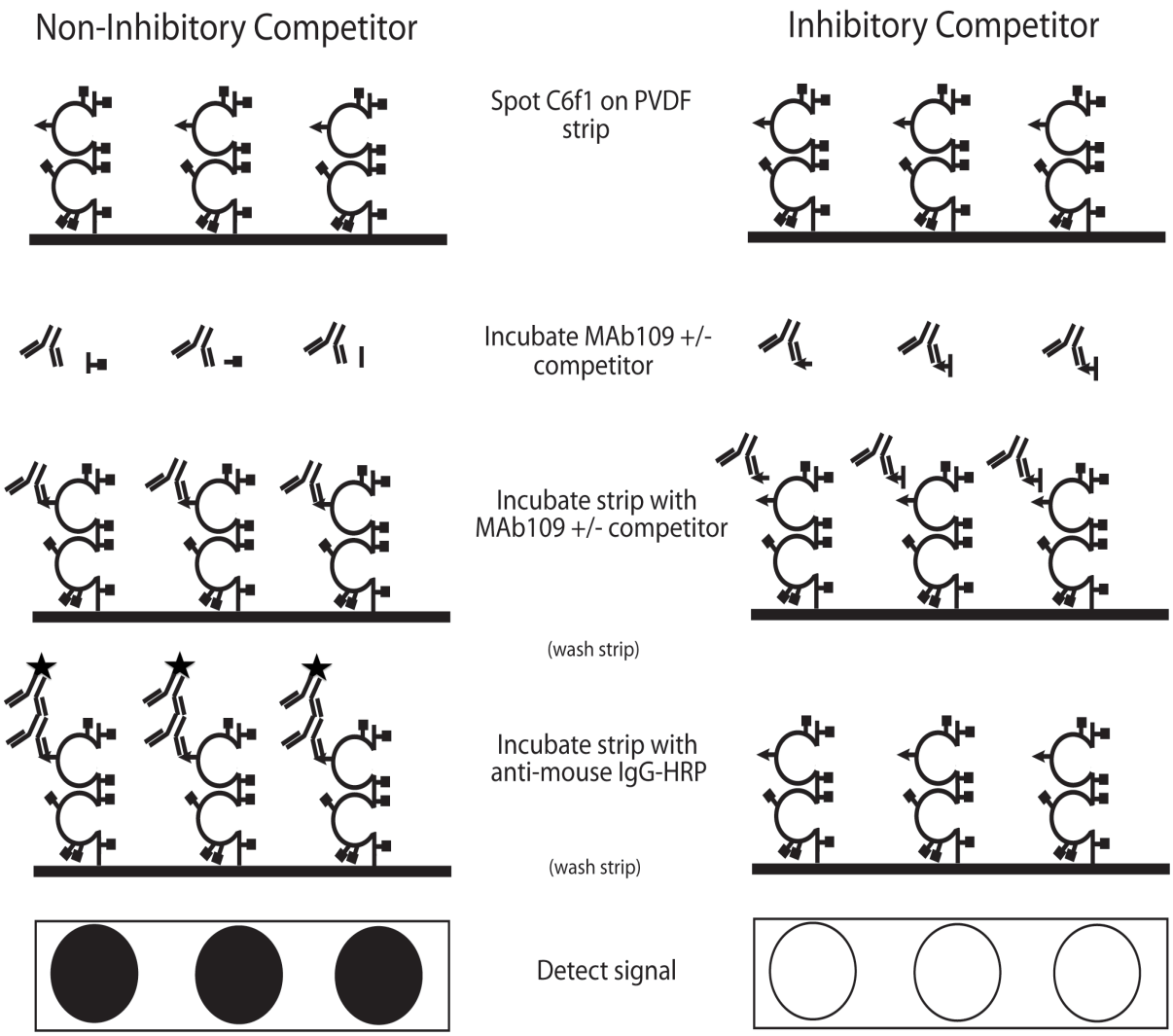


Figure 2.8. SCHEMATIC OF COMPETITIVE INHIBITION ASSAY. Purified C6f1 is spotted on PVDF strips in triplicate. C6f1 is depicted as two immunoglobulin domains with the N-linked glycosylation sites, which are either MAb109-reactive (triangle) or non-MAb109 reactive (square). MAb109 diluted in blocking buffer is incubated with each individual competitor for 30 minutes at room temperature. Competitors could be glycans (triangle or square only), glycopeptides (triangle or square with perpendicular line), or peptides (straight line only). This gives the antibody a change to bind the MAb109 epitope if it is present (right column); however, if the MAb109 epitope is not present then the antibody will not bind to the competitor tested (left column). After incubating the competitor with the antibody, the solution is added to the C6f1 strip. If the competitor did not contain the MAb109 epitope, MAb109 is free to bind to the MAb109 epitope found on the C6f1 spotted on the strip (left column). If the competitor did contain the MAb109 epitope, then MAb109 is already bound and cannot bind to the MAb109 found on C6f1 spotted on the strip (right column). Washing the membrane will remove the MAb109 bound to competitor since it is not bound to the C6f1 on the membrane. A secondary antibody, anti-mouse IgG-HRP (horse raddish peroxidase), is incubated with the strips to bind any MAb109 that may have bound to the C6f1 on the

strip. Washing removes any secondary antibody that did not bind to MAb109. The strips are then treated with enhanced chemiluminescent (ECL) and exposed to X-ray film to detect the signal from bound MAb109. Non-inhibitory competitors will give a signal (dark spots on X-ray film) since the MAb109 was able to bind to the MAb109 epitope found on the C6f1 on the strip. Inhibitory competitors will have reduced or absent signal (faint or no spots on X-ray film) depending on the degree of inhibition because MAb109 previously bound to the competitor, and therefore, cannot bind to the MAb109 epitope found on the C6f1 on the strip.

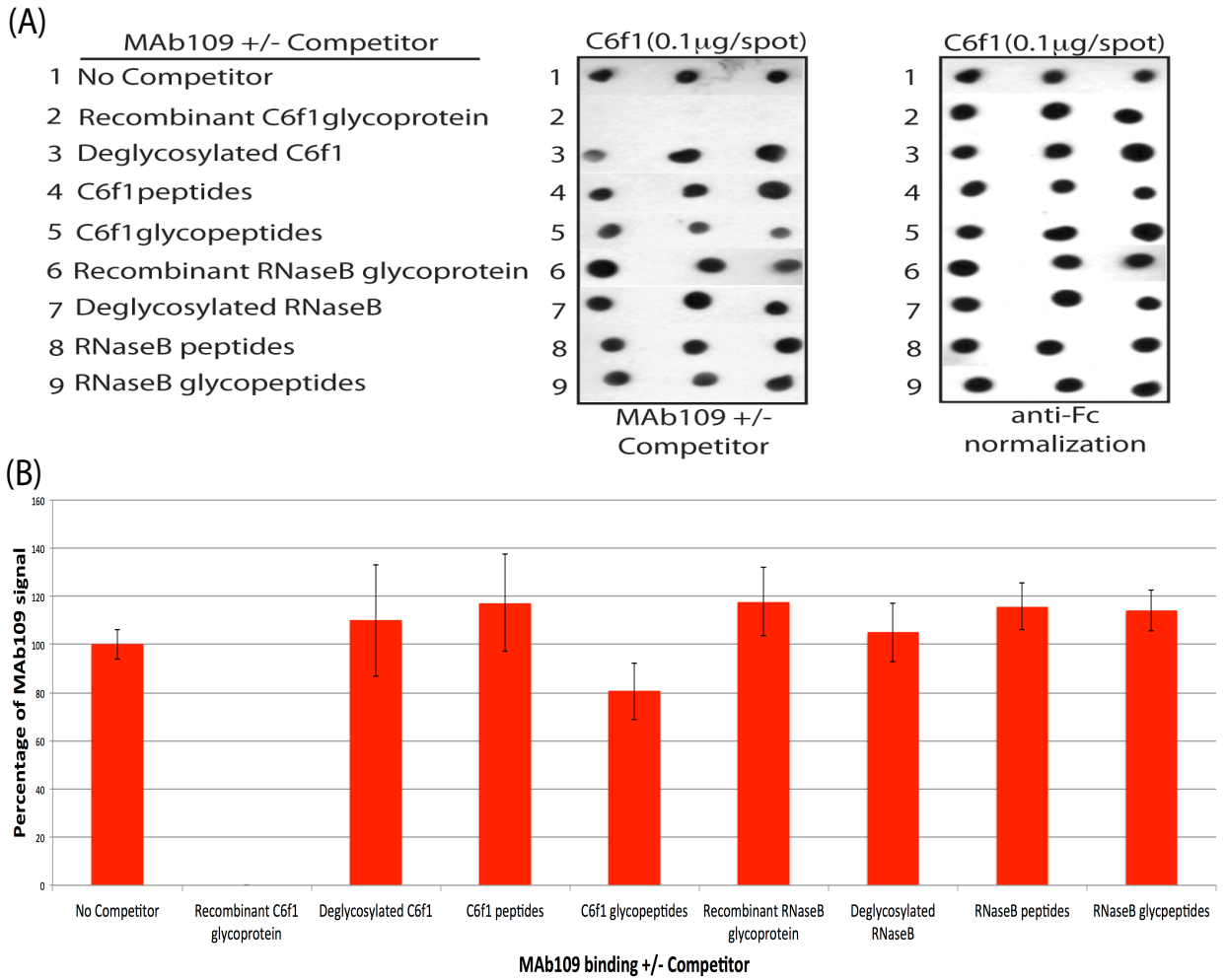


Figure 2.9. GLYCOPROTEINS AND GLYCANS DERIVED FROM PURIFIED C6F1 INHIBIT MAB109 BINDING. (A) Competitive inhibition assay using competitors derived from C6f1 expressed in HEK-293 cells and RNaseB. Upper left panel, competitors derived from C6f1 and RNaseB, a control glycoprotein, were incubated with MAb109 before immunoblotting on PVDF strips containing 0.1ug of C6f1 spotted in triplicate. Competitors in MAb109 binding assays were derived from C6f1 that was treated with various glycosidases and proteases, including PNGaseF, trypsin, and chymotrypsin, in order to produce glycopeptides, and peptides. To demonstrate that MAb109 inhibition was specific to C6f1 and its glycopeptides, RNaseB, a glycoprotein known to be N-linked glycosylated but not receive the MAb109 epitope, was used as a control and subjected to the same enzymatic treatments as C6f1. Shown in the center of panel A, the MAb109 +/- competitor immunoblot shows that recombinant C6f1 and C6f1 glycopeptides inhibit MAb109 binding while C6f1 protein or peptides did not. Additionally, no competitors derived from RNaseB inhibited MAb109 binding. Taken together, this suggests that N-link glycan(s) expressed on C6f1 are required for MAb109 inhibition. The upper right panel shows immunoblots of anti-Fc reactivity, which normalizes the amount of C6f1 spotted and demonstrates that the glycoprotein was present. (B) Graph quantifying the amount of inhibition. Inhibition of MAb109 binding +/- competitor to C6f1 was quantified using densitometry measurements taken with

ImageJ. Quantification of the ratio of MAb109 signal intensity to anti-Fc signal intensity for the dot blots found in panel A. The ratio for of MAb109 signal without competitor to its anti-Fc signal was normalized to 100% of MAb109 signal to which all other competitors were compared.

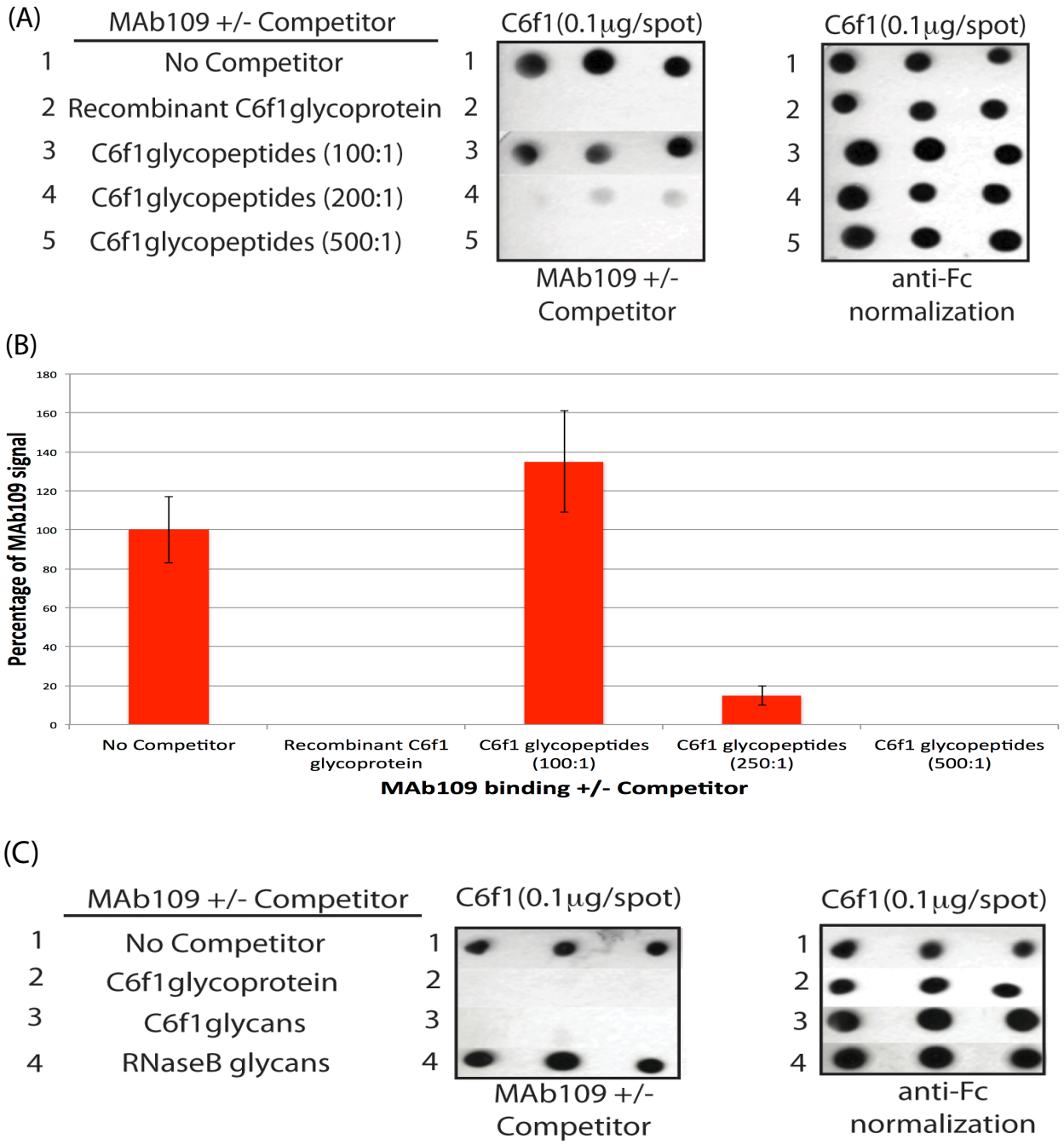
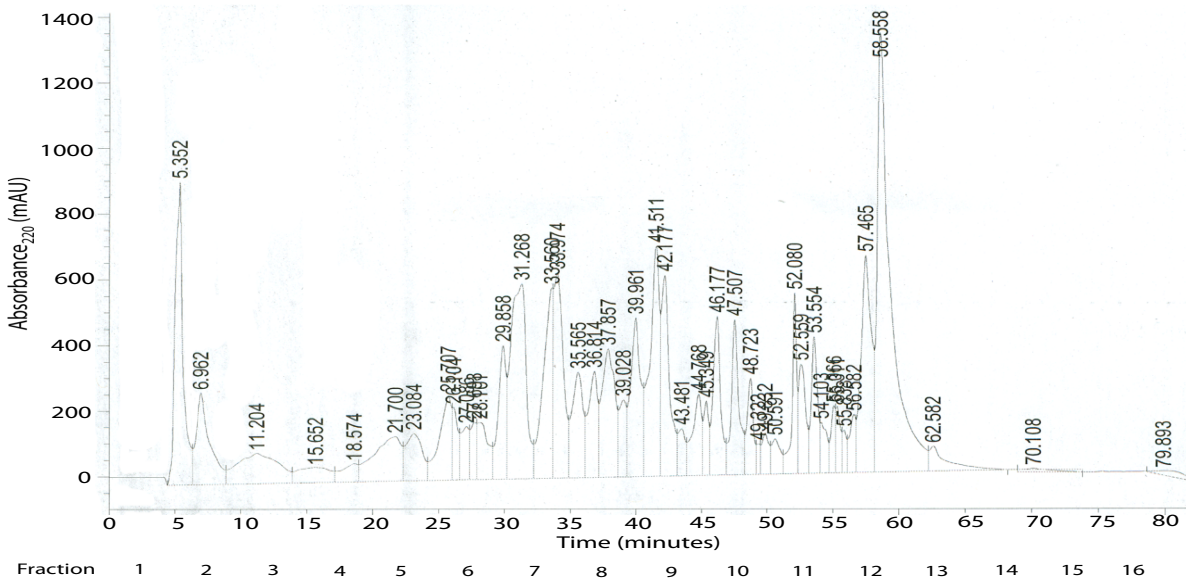


Figure 2.10. GLYCOSYLATED C6F1 GLYCOPROTEIN AND GLYCOPEPTIDES INHIBIT MAB109 BINDING. (A) Competitive inhibition assay using C6f1 glycopeptides results in decreased MAb109 binding in a concentration-dependent manner. Upper left panel shows that increasing the amount of C6f1 glycopeptides decreases the amount of MAb109 reactivity observed by immunoblot. Complete inhibition of the MAb109 signal results when C6f1 is in approximately 500 fold excess to MAb109 (lane 5); therefore, approximately 16 ng of C6f1 glycopeptides are required to completely inhibit the 0.3 ng of MAb109 used in the competitive inhibition assay. The upper right panel shows anti-Fc reactivity which was used to normalize the amount of C6f1 spotted on each strip. (B)

Graph quantifying the amount of inhibition by competitors incubated with MAb109. C6f1 intact glycoprotein and glycopeptides inhibit MAb109 reactivity. Inhibition of MAb109 binding +/- competitor was quantified using densitometry measurements taken with ImageJ. Quantification of the ratio of MAb109 signal intensity to anti-Fc signal intensity for the dot blots found in panel A. The ratio for of MAb109 signal without competitor was normalized to 100% of MAb109 signal to which all other competitors were compared. (C) Competitive inhibition assay using C6f1 glycans, but not RNaseB glycans results in decreased MAb109.

(A)


Absorbance220 (mAU)
Time (minutes)
Fraction
5.352
6.962
11.204
15.652
18.574
21.700
23.084
29.858
31.268
35.565
36.814
37.857
39.028
39.961
41.511
42.177
43.481
46.177
47.507
48.723
52.080
52.559
53.554
54.103
57.465
58.558
62.582
70.108
79.893

(B)

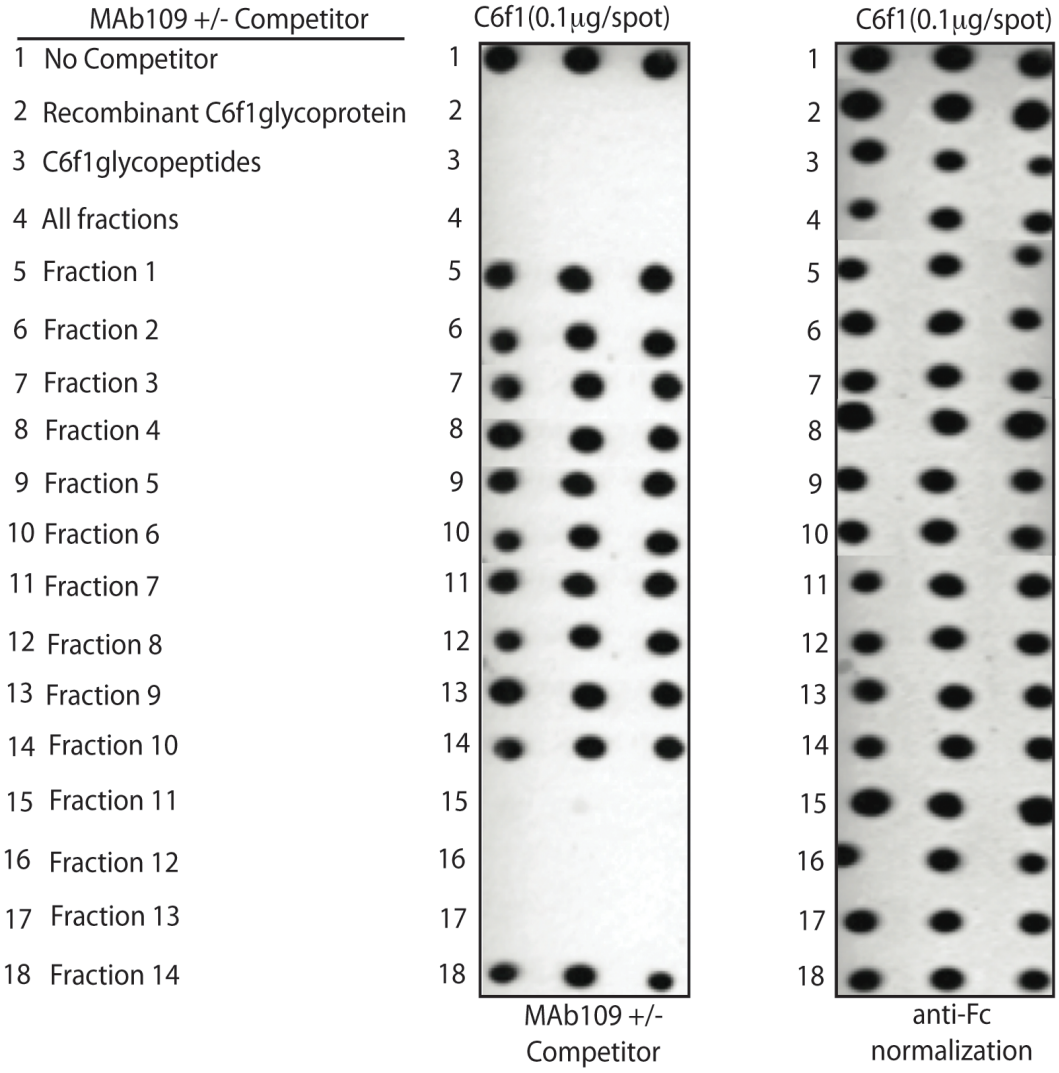
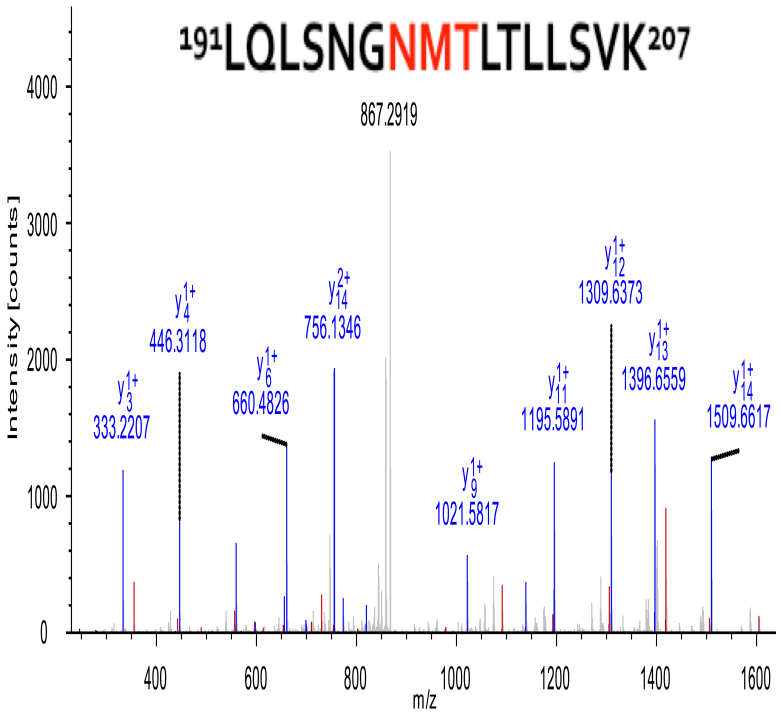

MAb109 +/- Competitor
1 No Competitor
2 Recombinant C6f1glycoprotein
3 C6f1glycopeptides
4 All fractions
5 Fraction 1
6 Fraction 2
7 Fraction 3
8 Fraction 4
9 Fraction 5
10 Fraction 6
11 Fraction 7
12 Fraction 8
13 Fraction 9
14 Fraction 10
15 Fraction 11
16 Fraction 12
17 Fraction 13
18 Fraction 14
C6f1(0.1μg/spot)
MAb109 +/- Competitor
C6f1(0.1μg/spot)
anti-Fc normalization

Figure 2.11. SPECIFIC C6F1 GLYCOPEPTIDES FRACTION INHIBIT MAB109 BINDING AFTER REVERSE PHASE CHROMATOGRAPHY. (A) Reverse phase chromatography trace at 220 nm of C6f1 glycopeptides separated using a C18 column. Dried tryptic/chymotryptic C6f1 glycopeptides were resuspended with 50 μL of buffer A and then separated by reverse phase liquid chromatography. Solvent A (0.1% trifluoroacetic acid (TFA)) and solvent B (80% acetonitrile, 0.085% TFA) were used to develop a linear gradient consisting of 5 min at 95% solvent A, 60 min gradient at variable slope to 95% solvent B, 3 min at 95% solvent B, 5 min to 95% solvent A, and 16 min at 100% solvent A. The glycopeptides were monitored at 220nm, and the trace is shown. Fractions were collected every 5 min and were dried down in a SpeedVac. (B) Competitive inhibition of MAb109 using glycopeptides fractionated by reverse phase chromatography. Previously dried, fractionated C6f1 glycopeptides were resuspended in water and 10% of each fraction was incubated separately with MAb109 before incubating with C6f1 spotted in triplicate on PVDF. The left side of panel B shows that fractions 11, 12, and 13 inhibited MAb109 reactivity signifying the presence of the MAb109 epitope in these fractions. The other fractions do not contain the MAb109 epitope as demonstrated by the lack of inhibition. Anti-Fc reactivity was used to normalize the amount of C6f1 spotted on the PVDF strips. The three fractions that inhibited MAb109 reactivity were pooled and dried down for peptide sequencing and glycan analysis.

(A)



(B)

#1	b(1+)	b(2+)	Seq.	y(1+)	y(2+)	#2
1	114.09135	57.54931	L			16
2	242.14993	121.57860	Q	1637.86652	819.43690	15
3	355.23400	178.12064	L	1509.80794	755.40761	14
4	442.26603	221.63665	S	1396.72387	698.86557	13
5	556.30896	278.65812	N	1309.69184	655.34956	12
6	613.33043	307.16885	G	1195.64891	598.32809	11
7	730.36162	365.68445	eamidated:18(	1138.62744	569.81736	10
8	877.39703	439.20215	M-Oxidation	1021.59625	511.30176	9
9	978.44471	489.72599	T	874.56083	437.78405	8
10	1091.52878	546.26803	L	773.51315	387.26021	7
11	1192.57646	596.79187	T	660.42908	330.71818	6
12	1305.66053	653.33390	L	559.38140	280.19434	5
13	1418.74460	709.87594	L	446.29733	223.65230	4
14	1505.77663	753.39195	S	333.21326	167.11027	3
15	1604.84505	802.92616	V	246.18123	123.59425	2
16			K	147.11281	74.06004	1

Figure 2.12. MAB109 EPITOPES IS FOUNDS ON A SINGLE, 16 AMINO ACID PEPTIDES OF C6F1. (A) MS spectrum of peptides isolated from fractions containing MAb109 reactive glycopeptides. Peptides were prepared by treating pooled, MAb109 reactive glycopeptide fractions with PNGaseF to release N-linked glycans. Peptides were separated from glycans by passing through a C_{18} column. Bound peptides were eluted using 40% isopropanol in 5% acetic acid and dried down. LC-MSMS of the peptides was performed with a program by which the 5 most intense peaks on each full mass spectrum were captured to be subjected to fragmentation (Top6) MS data was analyzed using Sequest requiring fully tryptic/ chymotryptic peptides and allowing for dynamic modification of Cys +57 (alkylation with iodoacetamide), Met +16 (oxidation), and Asn +3 (site of N-linked glycosylation) with Proteome Discoverer software (ThermoScientific). Peptide analysis identified a single 16 amino acid peptide corresponding to amino acids 191 to 207 of CEACAM6 sequence. Additionally, the identified peptide contained a predicted N-linked glycosylation sequon shown in red. (B) Table of m/z peaks corresponding the b- and y- ions for amino acids identified in the 16 amino acid peptide. The two left columns show the m/z values for b-ions in the 1+ and 2+ charge states while the two right columns show m/z values for y-ions in the 1+ and 2+ charge states. The values correspond to the amino acid sequence of the peptide shown in the center column and match the ^{191}LQLSNGNMTLTLLSVK^{207} sequence common to both C6f1 and CEACAM6.

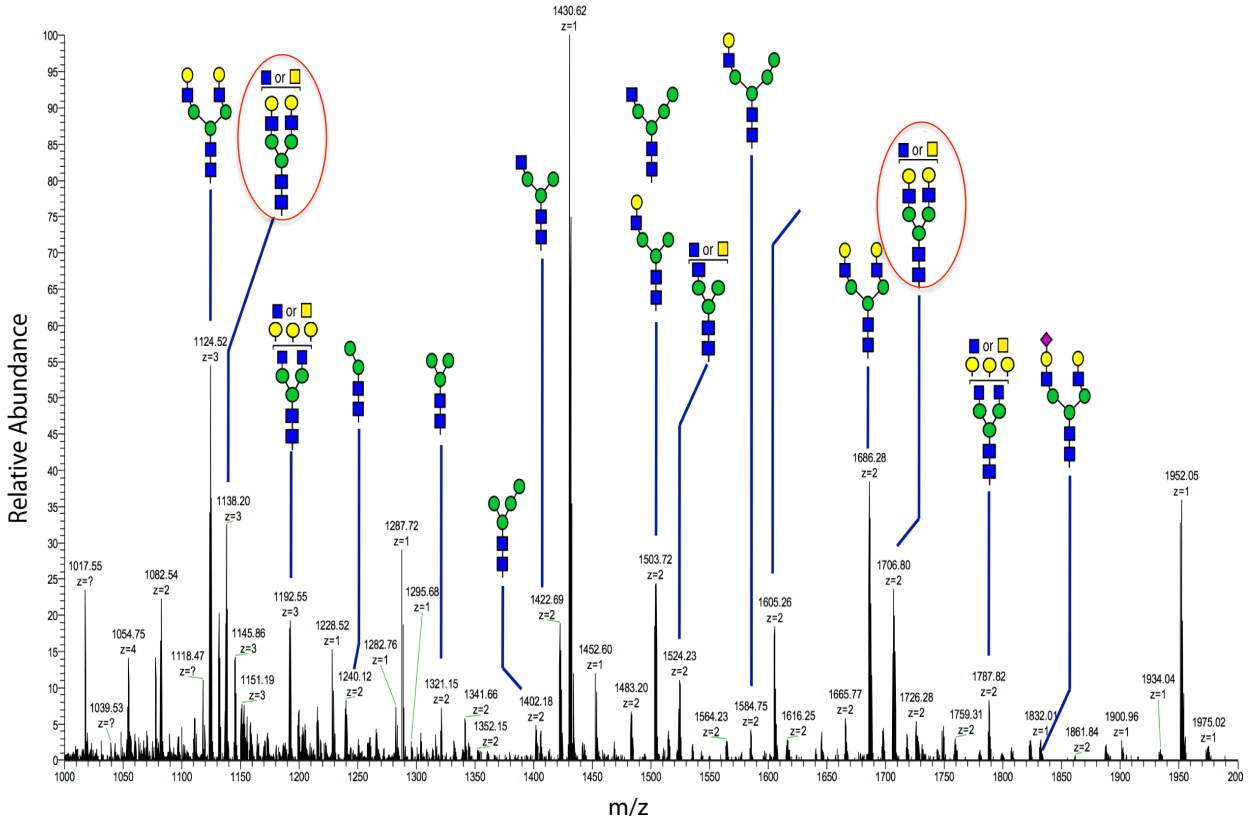


Figure 2.13. MOST PREVALENT N-LINKED MAB109-REACTIVE GLYCAN WAS A BIANTENNARY GLYCAN TERMINATING IN GALACTOSE CONTAINING AN ADDITIONAL HEXNAC. Total MS spectrum of permethylated N-linked glycans in purified MAb109-reactive glycopeptide fractions from HEK-293 cells. N-linked glycans were released from purified MAb109-reactive glycopeptide fractions. Purified N-linked glycans were permethylated and analyzed by nanospray mass spectrometry. The MS spectrum obtained by total ion monitoring shows that the most abundant MAb109 reactive glycan was a biantennary N-linked glycan terminating with galactose with an additional HexNAc, found in two mass-to-charge states (+2 and +3) (circled in red). The N-linked structures containing the additional HexNAc were annotated with an additional GlcNAc or GlcNAc over the glycan.

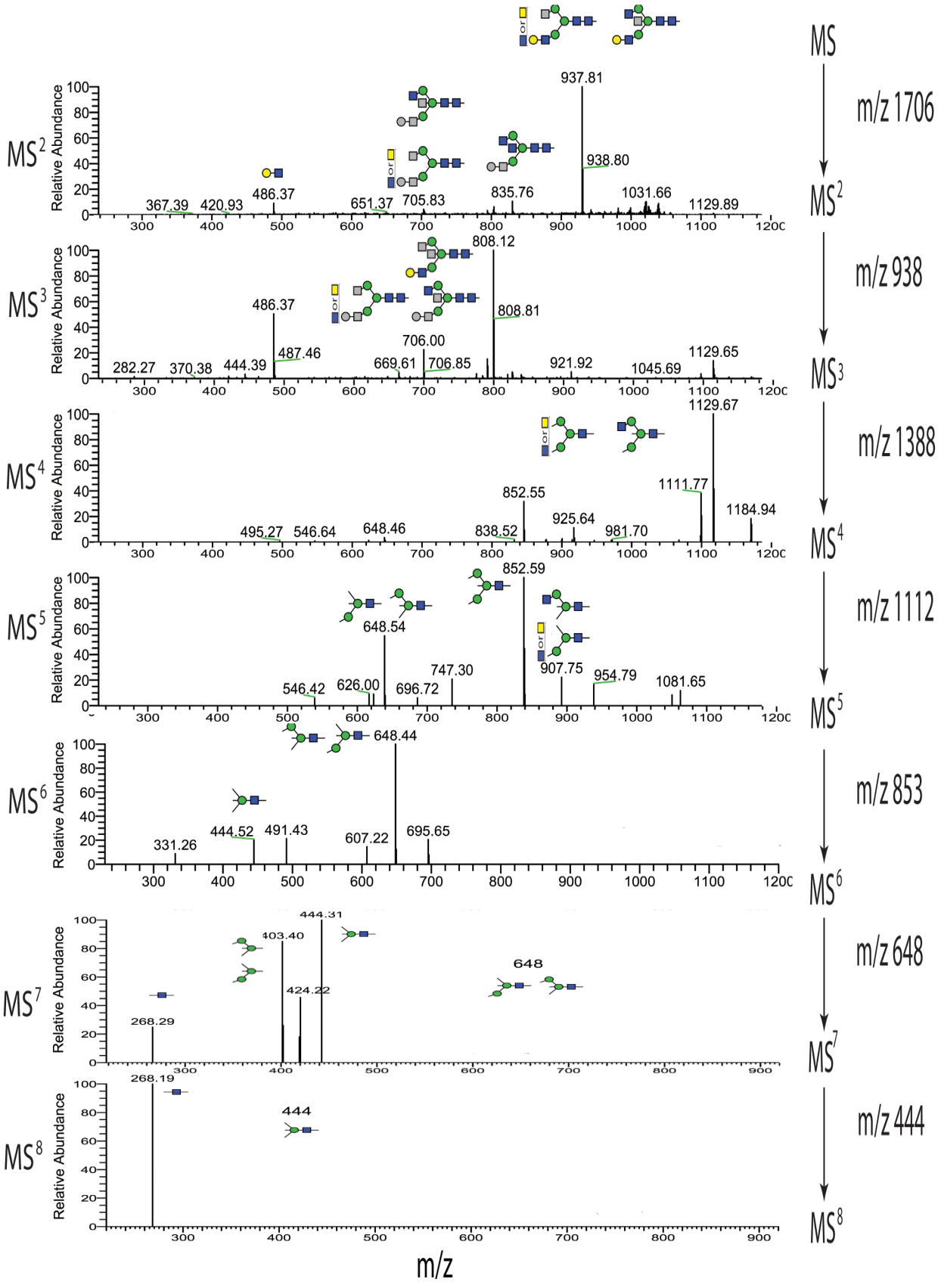

MS
m/z 1706
MS2
m/z 938
MS3
m/z 1388
MS4
m/z 1112
MS5
m/z 853
MS6
m/z 648
MS7
m/z 444
MS8
MS2
Relative Abundance
937.81
938.80
486.37
835.76
1031.66
1129.89
367.39
420.93
651.37
705.83
MS3
808.12
808.81
486.37
706.00
487.46
1129.65
282.27
370.38
444.39
669.61
706.85
921.92
1045.69
MS4
1129.67
852.55
1111.77
1184.94
495.27
546.64
648.46
838.52
925.64
981.70
MS5
852.59
648.54
907.75
747.30
954.79
1081.65
546.42
626.00
696.72
MS6
648.44
444.52
491.43
331.26
607.22
695.65
MS7
444.31
403.40
648
424.22
268.29
MS8
268.19
444
m/z

Figure 2.14. ADDITIONAL HEXNAC RESIDUE OF MAB109-REACTIVE GLYCAN IS FOUND ON THE β-LINKED MANNOSE OF THE TRISACCHARIDE CORE. Spectra were taken from the original m/z 1706 parent ion shown in Figure 13, a biantennary N-linked glycan terminating in galactose and had an additional HexNAc. The scheme for analyzing MS/MS spectra is outlined above on the right side. MS^2 shows peaks at m/z values of 938, 836, and 706 which are values that correspond to predicted mass glycans that have an additional HexNAc attached to the β-linked mannose. Successive MS^3 was performed on the most abundant 938 peak from MS^2, revealing fragments that corresponded to the HexNAc being attached to the β-linked mannose. The m/z peaks of 1112 in MS^4 and of 907 in MS^5 provide fragments with the HexNAc being attached to the β-linked mannose. MS^6, MS^7, and MS^8 are spectra showing fragmentation of the mannoses and second GlcNAc of the chitobiose core attached to the β-linked mannose, and the spectra were consistent with fragmentation expected if the HexNAc were attached to the β-linked mannose. m/z peaks that could have distinguished the HexNAc residue as GalNAc or GlcNAc were not observed.

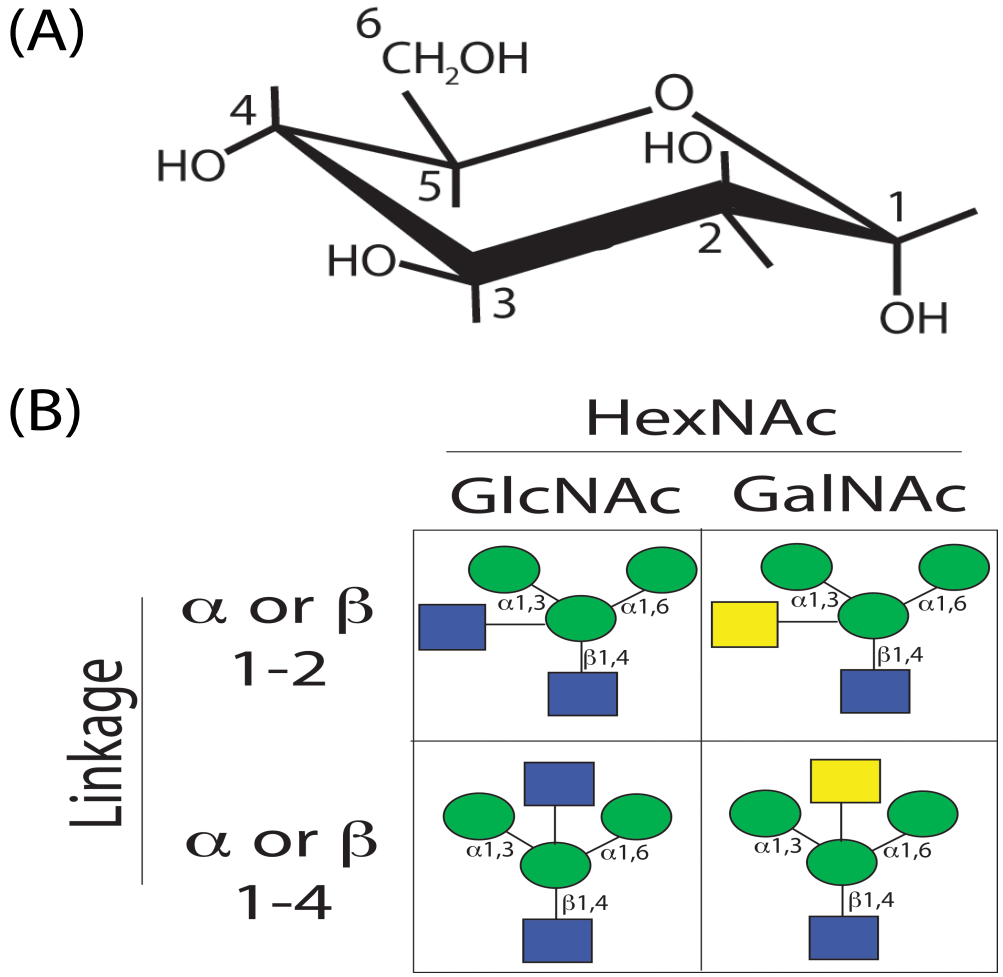


Figure 2.15. THE HEXNAC RESIDUE OF INTEREST COULD EITHER BE GALNAC OR GLCNAC BOUND AT EITHER THE 2- OR 4-POSITION OF THE β-LINKED MANNOSE. (A) Schematic of sugar residues attached to the β-linked mannose of a N-linked glycan. The GlcNAc from the chitobiose core is bound in the β-conformation at the 1-position, Man is bound in the α-confomation at the 3-position, and an additional Man is bound in the α-conformation at the 6 position. The only free hydroxyl groups available for the formation of additional glycosidic linkages are at the 2- and 4-positions. As a result, the additional HexNAc liked to the β-linked mannose identified by MS must be added at one of these two positions. (B) Illustration depicting the possible structures of an additional HexNAc bound to the β-linked mannose of a N-linked glycan. The possible structures exist because of the α and β anomers, two possible linkages,1-2 or 1-4, and two possible sugars, GlcNAc or GalNAc. C6f1 likely expresses one of these epitopes which is required for MAb109 binding.

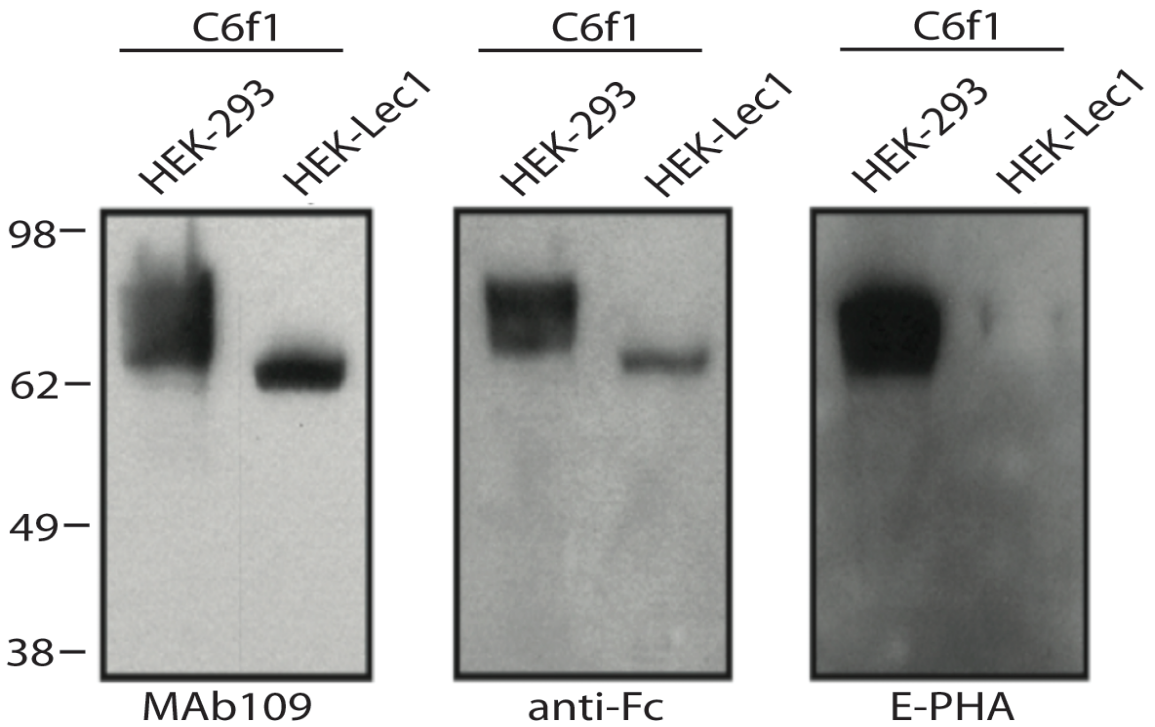


Figure 2.16. MAB109 EPITOPE EXPRESSED IN HEK LEC1 CELLS IS NEGATIVE FOR E-PHA REACTIVITY. The left, center, and right panels show the results from SDS-PAGE followed by MAb109 immunolotting, anti-Fc immunoblotting, and E-PHA lectin blotting, respectively, of supernatants taken from HEK wt type and HEK Lec1 cells transfected with C6f1. The left panel shows that C6f1 is expressed as a single, 85 kDa MAb109 reactive band in HEK-293 cells and as a single, 63 kDa MAb109 reactive band in HEK Lec1 cells. The spread in molecular weight observed for HEK wt when compared to HEK Lec1 suggest the possibility of more heterogeneous glycans in wt cells compared to Lec1 cells, which can only synthesize high-mannose glycans. The center panel shows the C6f1 expression level reflected by anti-Fc reactivity. Right panel shows E-PHA reactivity, which indicates the presence of bisected glycans on C6f1 expressed in HEK-293 cells but absence on C6f1 expressed in HEK Lec1 cells. The absence of E-PHA reactivity on C6f1 expressed in HEK-Lec1 cells suggests that MAb109 epitope is not due to the bisected (β1-4) GlcNAc.

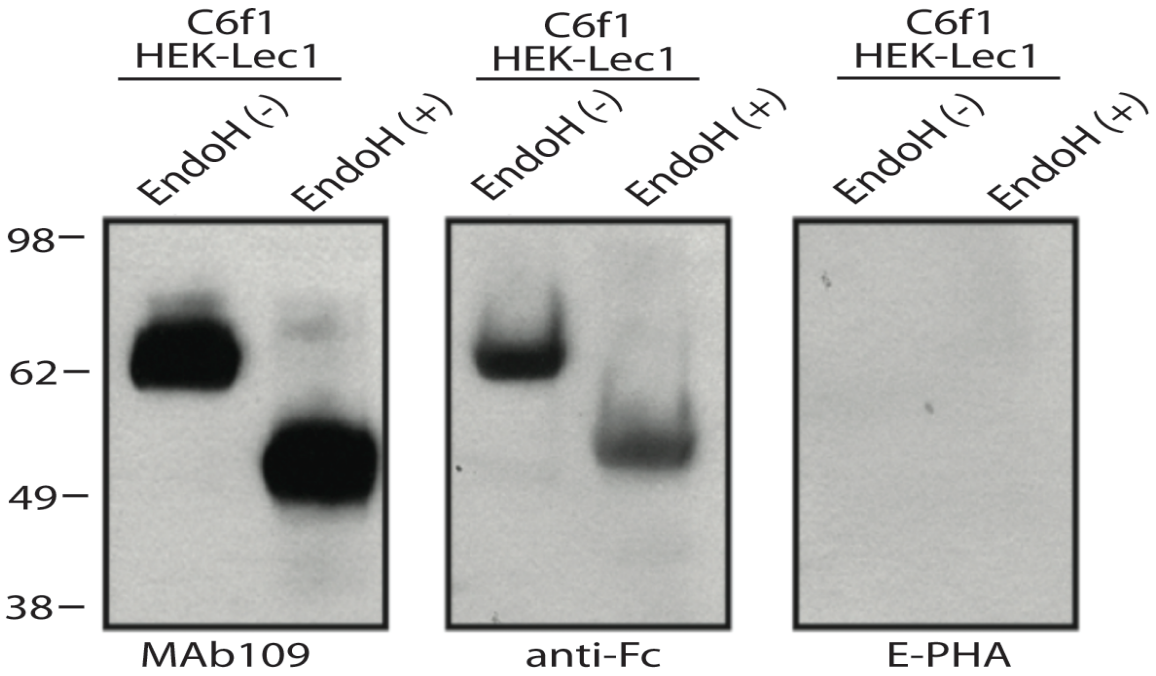


Figure 2.17. MAB109 EPITOPE EXPRESSED ON HIGH-MANNOSE GLYCAN OF HEK LEC1 CELLS IS ENDOH RESISTANT. The left, center, and right panels show the results of SDS-PAGE followed by MAb109 immunolotting, anti-Fc immunoblotting, and E-PHA lectin blotting, respectively, of supernatants taken from HEK Lec1 cells expressing C6f1 (+) or (-) EndoH treatment. The left panel shows that C6f1 expressed in HEK Lec1 cells remains MAb109-reactive after EndoH treatment, despite an approximately 13 kDa difference in molecular mass after EndoH treatement. The center panel shows that roughly equal amounts of glycoprotein were loaded as shown by anti-Fc reactivity. The right panel shows C6f1 expressed in HEK-Lec1 cells lacks E-PHA reactivity with or without EndoH treatment.

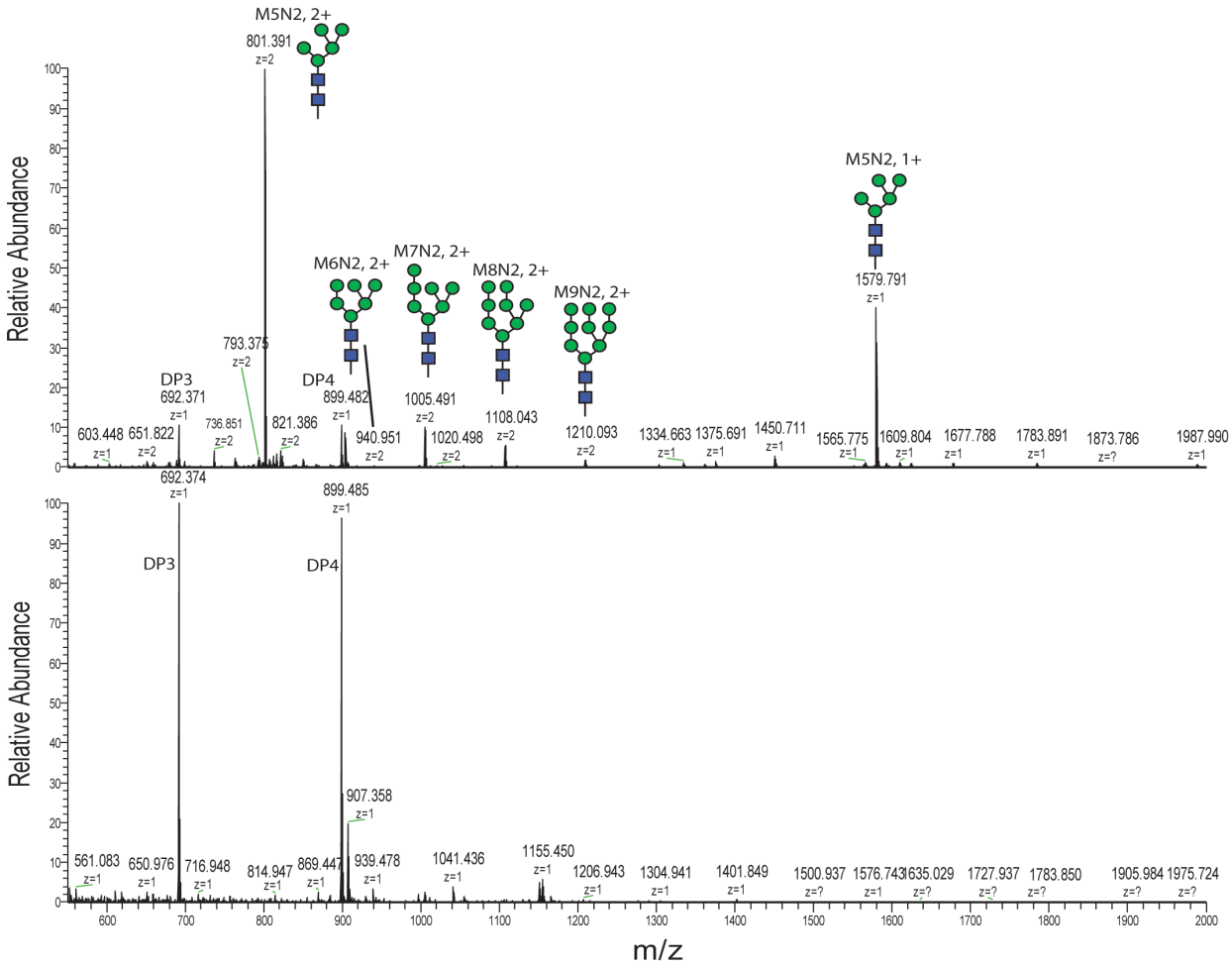


Figure 2.18. MS ANALYSIS OF HEK LEC1 EXPRESSED C6F1 GLYCANS. MS analysis of permethylated glycans after PNGaseF treatment of C6f1 expressed in HEK-LEC1 cells showed the presence of the high-mannose series of glycans (upper panel). Man_5- was the most prevalent glycan and was present at two different m/z (1+ and 2+). The signal intensity for Man_5Glc_2 (M5N2) (1+) was approximately ten fold the signal intensity of the standards 3-O-β-cellobiosyl-D-glucose (DP3) and 3-O-β-cellotriosyl-D-glucose (DP4). MS spectra of C6f1 glycans treated with EndoH before PNGaseF release, permethylation, and analysis showed that the peaks corresponding to Man_5-, Man_6-, Man_7-, Man_8-, and Man_9-$GlcNAc_2$ glycans were absent after EndoH treatment (Fig. 18, lower panel). The small peak at 907.39 may have correspond to a Man_5Hex_3 in the 2+ charge state(M5N3, 2+); however the peak was not abundant enough to perform MS^n.

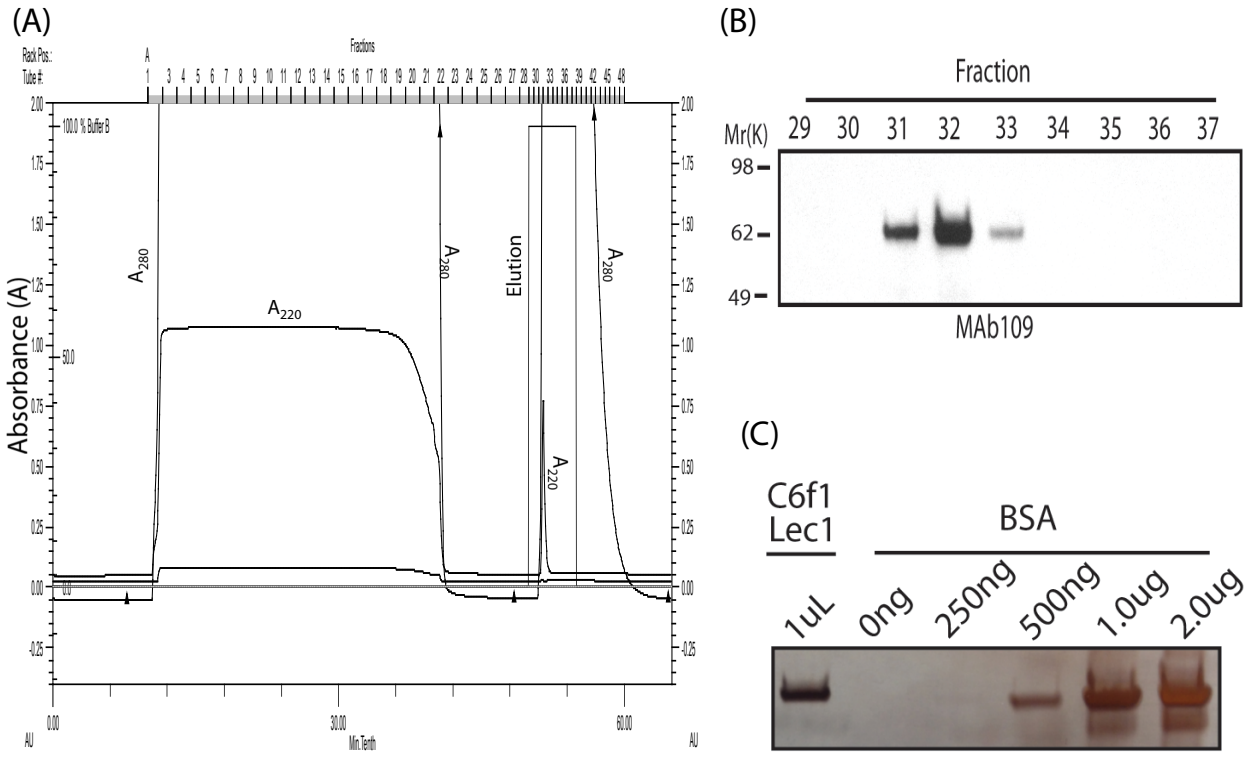


Figure 2.19. MILLIGRAM QUANTITIES OF RECOMBINANT C6F1 CAN BE PURIFIED USING TRANSIENTLY TRANSFECTED HEK-LEC1 SUSPENSION CELLS. (A) Chromatograph of supernatant taken from HEK-Lec1 cells transiently transfected with C6f1 and applied to immobilized protein A column. The column was washed before being eluted with 5 column volumes of 100mM citric acid pH 3.0. Elution of C6f1 was monitored by absorbance at 220 nm and 280 nm. The 280 nm trace is the solid line while the 220 trace is a solid line with small triangles. Fractions predicted to contain purified C6f1 were collected and immunoblotted with MAb109 following SDS-PAGE. (B) MAb109 immunoblot against fractions after low pH elution of protein A column. An approximately 63 kDa MAb109-reactive C6f1 band was present in the fractions corresponding to increased absorbance at 220nm (C) Silver stain of C6f1 expressed in HEK-Lec1 cells and purified using immobilized protein A. The amount of purified C6f1 was calculated by comparing C6f1 staining relative to BSA standards on the same gel. The yield of C6f1 was approximately 7.5 μg/mL of cell supernatant.

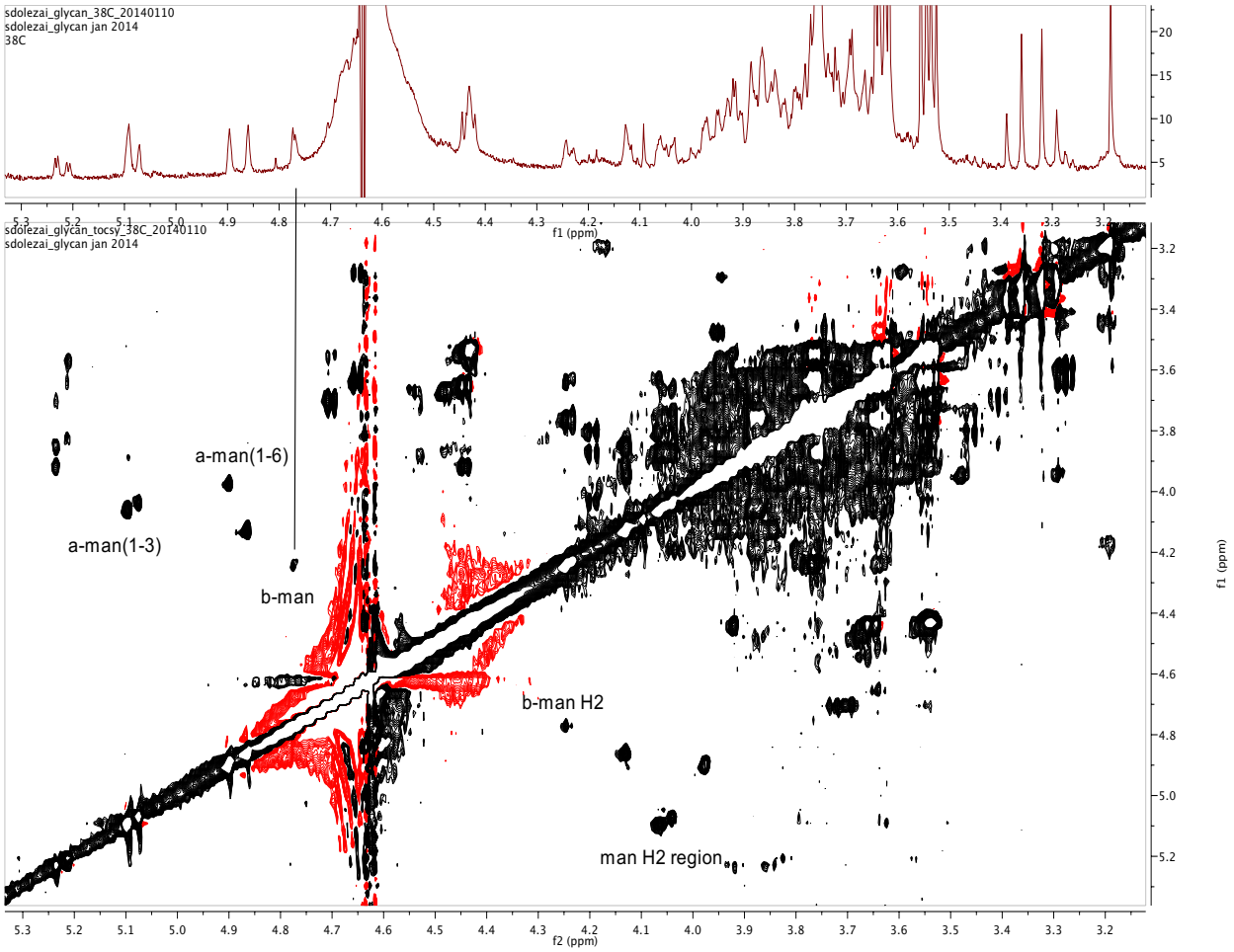


Figure 2.20. NMR ANALYSIS OF MAB109 REACTIVE GLYCAN ISOLATED FROM HEK-LEC1 EXPRESSED C6F1. Both one-dimensional and two-dimensional spectra for the MAb109-reactive glycans isolated from HEK-Lec1 expressed C6f1 were obtained using a 600 MHz instrument. In the one-dimensional spectra (upper panel), many peaks were observed between 4.0 and 5.5 ppm, an area where carbohydrates are commonly detected. α-linked and b-linked mannoses corresponding to a high-mannose structure were detected and annotated in the two-dimensional spectrum (lower panel) Additionally, the H2 cross peaks were observed in the two-dimensional spectra. Although, the mannoses of the glycan expressing the MAb109 were detected, the HexNAc could not be definitively identified. A peak at approximately 4.45 shows characteristics of both GalNAc and GlcNAc.

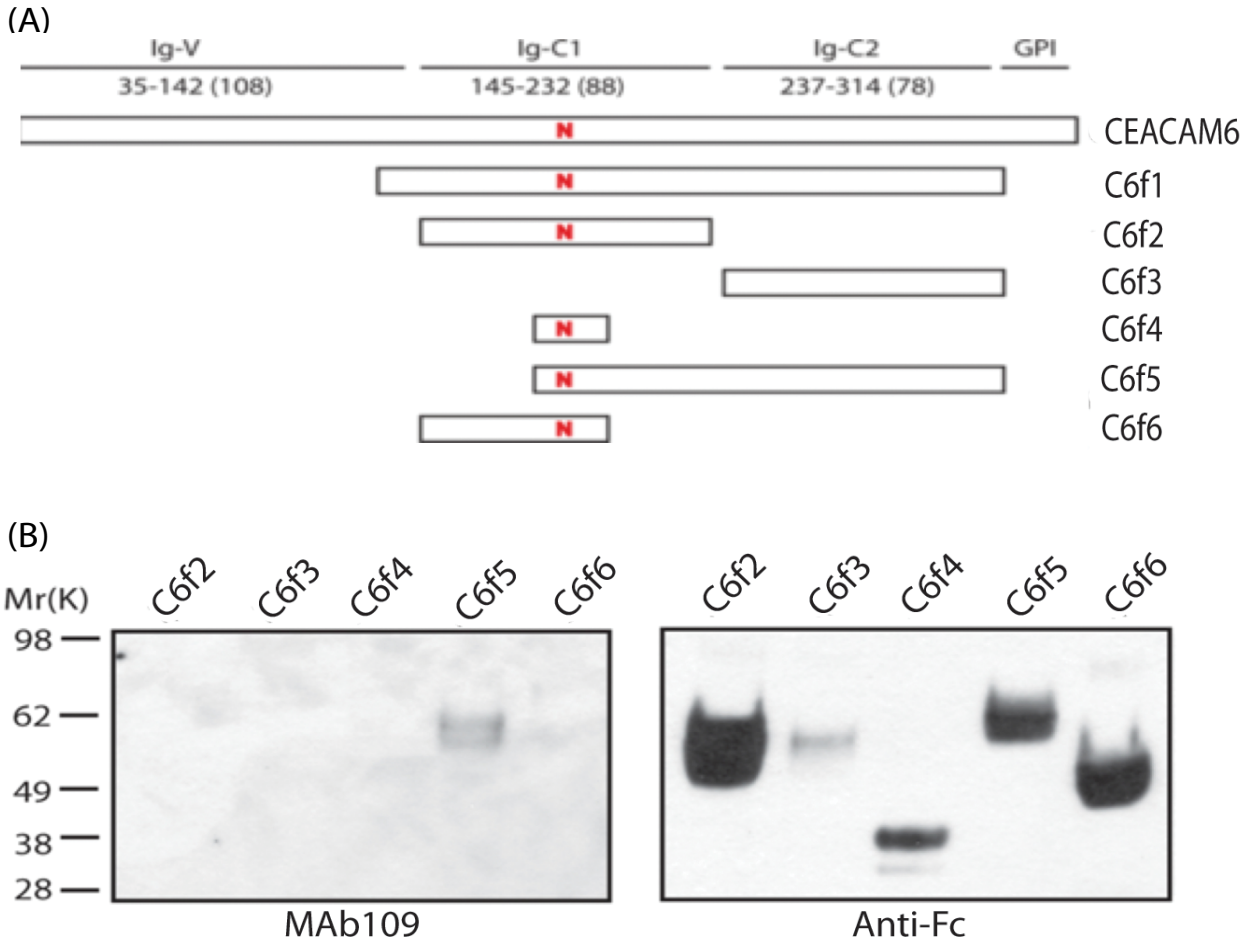


Figure 2.21. INVESTIGATION OF THE MINIMAL SEQUENCE OF C6F1 REQUIRED FOR MAB109 EPITOPE EXPRESSION. (A) Schematic showing various CEACAM6 fragments (C6f1-C6f6) in spatial relation to the full-length CEACAM6. CEACAM6 consists of three immunoglobulin domains and a GPI-anchor shown in relation to the full-length CEACAM6. Truncated cDNA-encoding fragments common to CEACAM6 and C6f1 were PCR amplified and ligated into the pFUSE vector. (B) Only the C6f5 fragment, which encodes for the MAb109 occupied N-linked glycosylation site and the downstream sequence of CEACAM6 was positive for MAb109 reactivity. Left panel shows that the C6f5 fragment was an approximately 58 kDa band after SDS-PAGE and was positive for MAb109 reactivity on immunoblot. The right panel of Figure 21B shows anti-Fc reactivity for all constructs demonstrating that all the constructs were expressed.

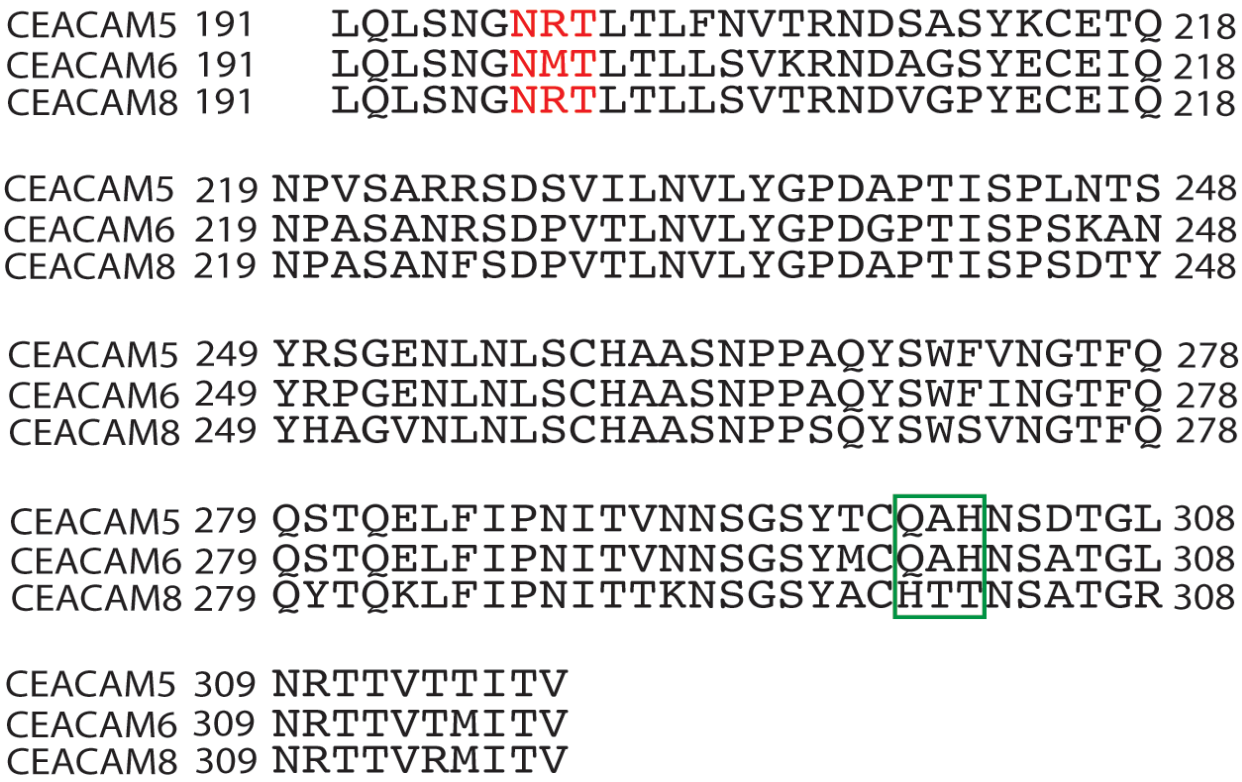


Figure 2.22. SEQUENCE ALIGNMENT OF THE C-TERMINI OF CEACAM5, CEACAM6, AND CEACAM8 IDENTIFIED A CLUSTER OF THREE AMINO ACIDS THAT VARIED BETWEEN MAB109 EPITOPE EXPRESSING AND NON-EXPRESSING GLYCOPROTEINS. Sequence alignment of MAb109 reactive proteins (CEACAM5 and CEACAM6) against a non-MAb109 reactive protein (CEACAM8) identified a three amino acid group, 300 to 302, common to MAb109-reactive glycoproteins that was clearly different in a non-reactive glycoprotein (boxed in green). These amino acids are found downstream from the glycosylation site, Asn-197, shown in red text, which expresses the MAb109 epitope.

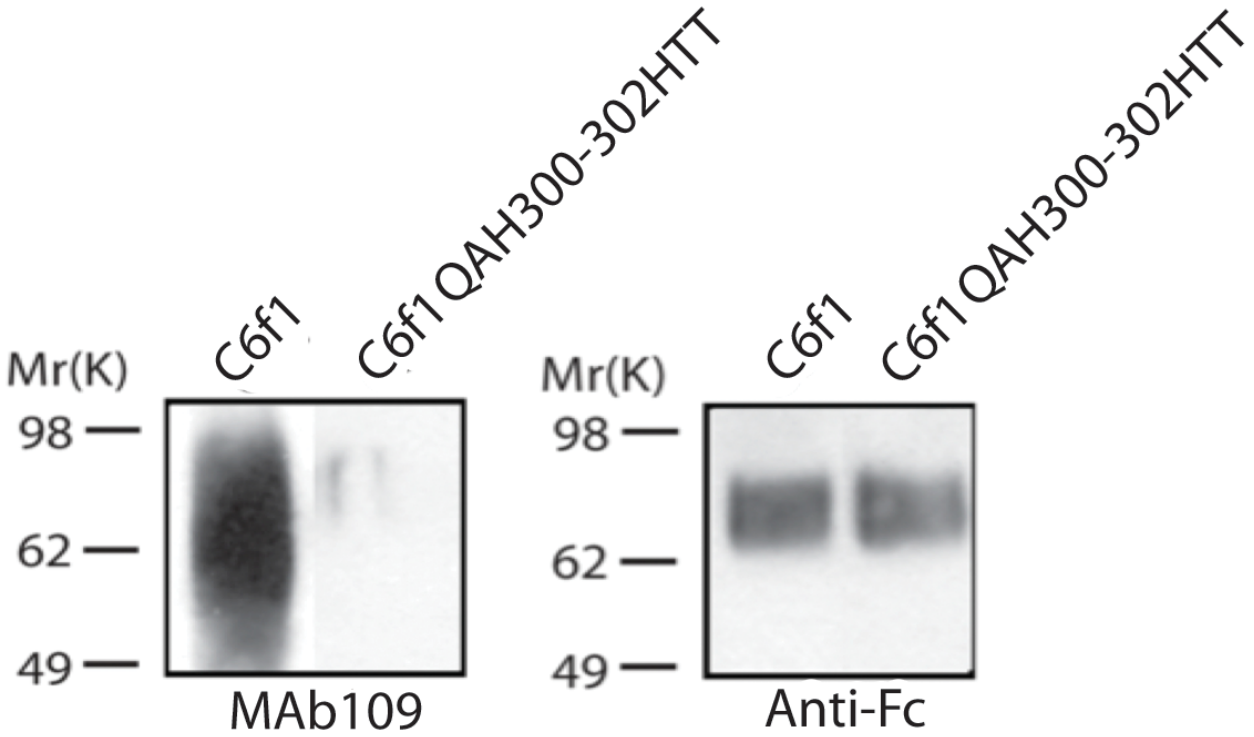


Figure 2.23. MUTATION OF AMINO ACIDS 300 TO 302 FROM CEACAM6 SEQUENCE TO THE CORRESPONDING AMINO ACIDS IN CEACAM8 ABOLISHES MAB109 EPITOPE EXPRESSION. Left panel**, i**mmunoblot of site directed mutant construct shows that changing C-terminus amino acids of CEACAM6, QAH, to that of non-MAb109 expressing CEACAM8, HTT, abolishes MAb109 reactivity. The MAb109 reactive C6f1 band encompassed a wider range of molecular weights due to the overexposure of the membrane to ensure the C6f1 mutant was negative for MAb109 reactivity. Right panel, immunoblot with anti-Fc antibody using supernatants collected from transfected HEK-293 cells showing that both C6f1 and the C6f1 mutant were expressed in equivalent amounts.

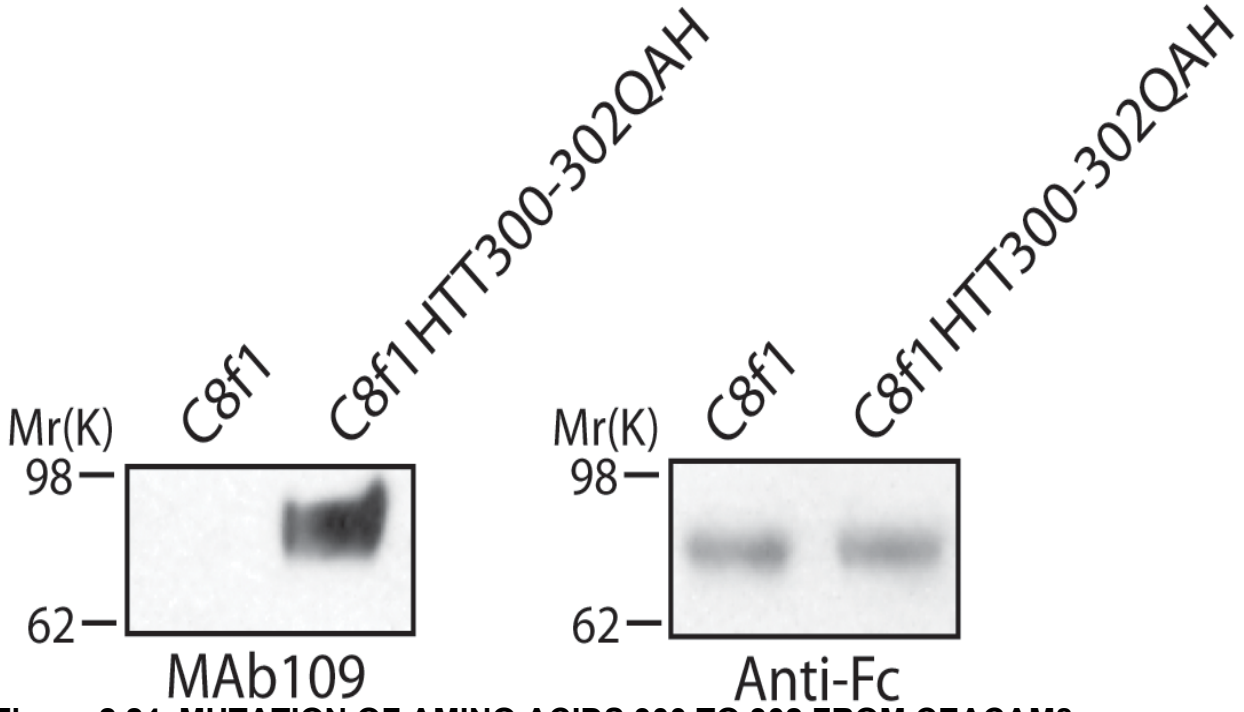


Figure 2.24. MUTATION OF AMINO ACIDS 300 TO 302 FROM CEACAM8 SEQUENCE TO THE CORRESPONDING AMINO ACIDS IN CEACAM6 RESULTS IN MAB109 EPITOPE EXPRESSION. Left panel, CEACAM8 fragment 1 (C8f1), a fragment overlapping the same number and position of amino acids as C6f1, did not express the MAb109 epitope when transfected into and secreted from HEK-293 cells. However, site-directed mutagenesis of amino acids 300 to 302 of C8f1 replacing the corresponding CEACAM8 sequence, HTT, with that of CEACAM6, QAH, resulted in MAb109 epitope expression. Right panel, immunoblot with anti-Fc antibody of supernatants collected from transfected HEK-293 cells showing that both C8f1 and the C8f1 mutant were expressed at equivalent levels.

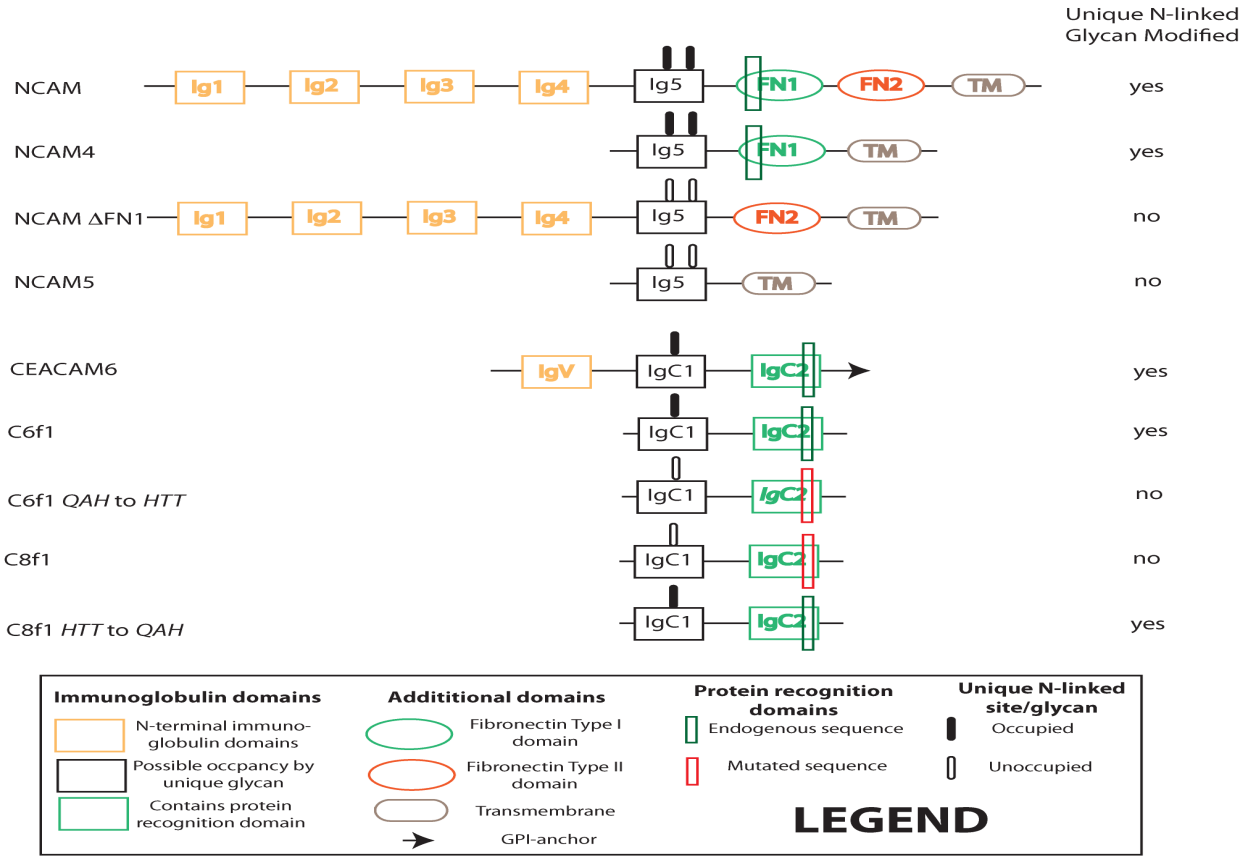


Figure 2.25. DIAGRAM ILLUSTRATING THE COMMONALITIES BETWEEN THE PROTEIN-SPECIFIC N-LINKED MODIFICATION OBSERVED IN THE POLYSIALYLATION OF NCAM AND THE MAB109 EPITOPE EXPRESSION ON CEACAM6. In both instances, an amino acid downstream of the occupied N-linked site is required for the addition of a unique glycan on an adjacent immunoglobulin domain. Removal or mutation of specific, C-terminus amino acid sequences results in the absence of the unique, N-linked glycan.

CHAPTER 3. APPLICATIONS AND FUTURE PERSPECTIVES

Pancreatic cancer is a deadly disease which is difficult to diagnose and for which the only curative treatment is surgery after early detection. Carbohydrate antigen 19-9 (CA19-9) has been the most commonly used pancreatic marker since it was described by Koprowski, in 1979 [45], but the test's low sensitivity and specificity has had limited improving patient outcomes over the last 25 years [46, 47]. Despite the limited success of CA19-9, glycans and glycoproteins are becoming increasingly recognized as potential cancer markers because glycosylation changes often occur during the early stages of oncogenesis, and they also provide a cell surface molecule that can be targeted by monoclonal antibodies. The investigation of glycans as biomarkers and the generation of monoclonal antibodies against them provide the potential means for improving the diagnosis and treatment of pancreatic cancer.

The use of one such glycan-specific antibody, MAb109, in this dissertation work has led to the identification of a novel glycan as a potential marker for pancreatic cancer. Previous immunohistochemical (IHC) staining with MAb109 showed that the antibody detected dysplastic pancreas and pancreatic cancer from patient biopsies. Staining additional pancreatic cancer sections as well as control sections from non-diseased pancreas, pancreatitis, and pancreatic precursor lesions such as intraductal papillary mucinous neoplasia (IPMN) and mucinous cystic neoplasia (MCN) will be important for determining the antibody's sensitivity and specificity. This proposed MAb109 IHC study and other studies including investigating the internalization of the antibody, determining

the antibody half-life, and measuring antibody-binding affinity will all contribute to predicting the effectiveness of MAb109 as a clinical tool in the detection and treatment of pancreatic cancer. In addition to its use as a diagnostic, the potential development of MAb109 into a monoclonal-based therapeutic is especially exciting. The development of biological drugs is becoming increasingly more prevalent, and MAb109 could provide a platform to develop a drug specifically against pancreatic cancer. The use of MAb109 in the early detection and treatment of pancreatic cancer is exciting considering the antibody has potential to improve patients' lives which is the primary goal of clinical research.

Although the findings from this dissertation have clinical application, the discovery of a novel, N-linked glycan that is specifically expressed on CEACAM6 will have importance in basic science as well. The identification of a specific cluster of amino acids required for addition of the novel, MAb109 epitope to CEACAM6 provides evidence that this modification is another of a few select instances of protein-specific glycosylation. The identification of the glycosyltransferase responsible for addition of the MAb109 epitope will be crucial moving forward. Since the MAb109 epitope is a novel structure, not previously described, and requires underlying protein recognition for its addition, the epitope is likely added by a previously unassigned open reading frame. Identifying the enzyme would allow the regulation of epitope expression and determine its functional role in cell adhesion, cell migration, and tissue architecture, which are affected by altered CEACAM6 expression [28, 30, 31]. Future experiments investigating the function of the epitope on CEACAM6 will be interesting in determining the effect that

altered MAb109 epitope and CEACAM6 expression have independently and cooperatively on cellular function and phenotype.

Although the function of the epitope is not currently known, the specificity of the modification and its expression in numerous cell lines suggest an underlying, important function. Why would a cell use unnecessary energy expenditure to synthesize a glycosyltransferase that adds to a few select acceptors that the cell may or may not see if the function of that addition was not important? Addition of the epitope may provide a means to fine tune the function of CEACAM6. Since CEACAM6 is an oncofetal antigen, epitope expression may decrease cell adhesions during gastrointestinal tract development favoring cell migration and plasticity during development. This state would also favored by tumors during oncogenesis as they are growing and invading into the surrounding non-diseased tissue. The study of the promoter regions of the glycosyltransferase responsible for MAb109 addition and its acceptors will be important in understanding the regulation of their expression. Additionally, real-time PCR studies of the glycosyltransferase adding the MAb109 epitope will provide evidence of whether the enzyme is expressed at all times or that it's expression is someway induced by the expression of its acceptor.

One of the most important findings of this dissertation was the identification of amino acid sequence determinants in CEACAM6 necessary for addition of the MAb109 epitope. For the majority of glycosyltransferases, acceptor recognition does not require the underlying polypeptides sequence; however, MAb109 epitope addition required specific amino acids in CEACAM6. These amino acids possibly allow the specific recognition of CEACAM6 by the glycosyltransferase responsible for adding the MAb109

epitope. The site specificity of MAb109 occupancy could possibly be explained by the resulting proximity of the N-linked site to the glycosyltransferase active site after enzyme binding to CEACAM6. Even though the MAb109 epitope prefers a specific N-linked site, the possibility that the modification could occur on another N-link sites cannot be ruled out and needs to be investigated.

Many more examples of protein-specific glycosylation are likely waiting to be discovered. Numerous sequences in the CAZy database are uncharacterized, and these could be enzymes which glycosylate specific glycans expressed on specific proteins and specific glycosylation sites on those proteins. Examples of protein-specific glycosylation are difficult to detect without the specificity provided by monoclonal antibodies. The MAb109 epitope would have been extremely difficult, if not impossible, to detect on one of twelve N-linked sites of one protein, CEACAM6, without the use of MAb109. Identification is further complicated when the modification is similar to a more prevalent, existing glycan as was the case with trying to differentiate the MAb109 epitope from the β1-4 GlcNAc of GnT-III. Identifying examples of protein specific glycosylation will improve as researchers develop better scientific tools and more sensitive instruments with which to detect novel glycan structures.

The MAb109 epitope was originally observed on CEACAM6; however, the epitope is expressed on proteins not yet identified. Unpublished evidence supports this hypothesis since MAb109-reactive bands have been observed in Drosophila, which does not express CEACAM5 or CEACAM6. They do, however, express members of the immunoglobulin (Ig) family of proteins [42]. Because of the identification of the polysialylation of NCAM and now the MAb109 epitope on CEACAM6, and Ig family of

proteins may supply additional examples of protein-specific glycosylation. The Ig-family is an evolutionally conserved protein structure that is expressed in numerous organisms. Although the Ig-fold is conserved, protein-specific glycosylation would provide a means to regulate the temporal, spatial, or functional characteristics of the protein on which it is found.

Altogether this dissertation has made contributions to the identification of a novel, N-linked glycan that is expressed by a specific glycoprotein in pancreatic cancer and will hopefully lead to the development of diagnostics and therapeutics for pancreatic cancer.

REFERENCES

1. Koprowski, H., et al., *Colorectal carcinoma antigens detected by hybridoma antibodies.* Somatic Cell Genet, 1979. **5**(6): p. 957-71.
2. Goonetilleke, K.S. and A.K. Siriwardena, *Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer.* Eur J Surg Oncol, 2007. **33**(3): p. 266-70.
3. Ballehaninna, U.K. and R.S. Chamberlain, *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. **3**(2): p. 105-19.
4. Blumenthal, R.D., H.J. Hansen, and D.M. Goldenberg, *Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen).* Cancer Res, 2005. **65**(19): p. 8809-17.
5. Duxbury, M.S., et al., *CEACAM6 is a determinant of pancreatic adenocarcinoma cellular invasiveness.* Br J Cancer, 2004. **91**(7): p. 1384-90.
6. Duxbury, M.S., et al., *CEACAM6 gene silencing impairs anoikis resistance and in vivo metastatic ability of pancreatic adenocarcinoma cells.* Oncogene, 2004. **23**(2): p. 465-73.
7. Vogel, C., S.A. Teichmann, and C. Chothia, *The immunoglobulin superfamily in Drosophila melanogaster and Caenorhabditis elegans and the evolution of complexity.* Development, 2003. **130**(25): p. 6317-28.

SUPPLEMENT: ELEVATED LEVELS OF GLYCOSYLPHOSPHATIDYLINOSITOL (GPI) ANCHORED PROTEINS IN PLASMA FROM HUMAN CANCERS DETECTED BY *C. SEPTICUM* ALPHA TOXIN

Samuel Dolezal, Shanterian Hester, Pamela S. Kirby, Allison Nairn, Michael Pierce, and Karen L. Abbott. Cancer Biomarkers. 2014 Jan 1;14(1):55-62. Reprinted here with permission of publisher.

ABSTRACT

The glycosylphosphatidylinositol (GPI) anchor is a glycan and lipid posttranslational modification added to proteins in the endoplasmic reticulum. Certain enzymes within the GPI biosynthetic pathway, particularly the subunits of the GPI transamidase, are elevated in various human cancers. Specific GPI anchored proteins, such as carcinoembryonic antigen and mesothelin, have been described as potential biomarkers for certain cancers; however, the overall levels of GPI anchored proteins present in serum from cases of human cancers have not been evaluated. We have developed the use of a bacterial toxin known as alpha toxin from *Clostridium septicum* to detect GPI anchored proteins *in vitro*. In this study, we use alpha toxin to detect GPI anchored proteins present in plasma from cases of several types of human cancers. Our data indicate that human cancers with previously documented elevations of GPI transamidase subunits show increased alpha toxin binding to plasma from patients with these cancers, indicating increased levels of GPI anchored proteins. Furthermore, our results reveal very low levels of alpha toxin binding to plasma from patients with no malignant disease indicating few GPI anchored proteins are present. These data suggest that GPI anchored proteins present in plasma from these cancers represent biomarkers with potential use for cancer detection.

INTRODUCTION

A particular glycosylation known as the glycosylphosphatidylinositol (GPI) anchor is a unique type of glycoconjugate added to certain cell surface glycoproteins that contain a C-terminal signal sequence in addition to the N-terminal signal sequence. This specialized glycan and lipid-containing membrane linkage is highly conserved in all eukaryotic species establishing the importance of this form of glycosylation. The GPI anchor consists of a core structure that is assembled and added en bloc to the C-terminus of proteins by the GPI transamidase (GPIT) multisubunit complex. Catalytic and non-catalytic subunits of the GPIT exhibit increased expression at the mRNA and protein level in many human cancers [48-50]. For example, breast carcinoma has significant elevations in the *GPAA1* and *PIGT* subunits of the GPIT due to chromosomal amplifications. The increased expression of these non-catalytic subunits was found to significantly increase both the tumorigenicity and the overall levels of GPI anchored proteins in breast carcinoma [48, 51].

The overall abundance of GPI anchored proteins is estimated to comprise 1-2 % of all translated proteins in the human proteome [52]. Their abundance in specific tissues may be lower due to the regulation of specifically expressed GPI acceptor proteins. We have been utilizing alpha toxin to capture and identify GPI anchored proteins by mass spectrometry [51]. Our studies have revealed that alpha toxin binds with the GPI glycan region as shown by retained binding of the toxin after removal of the lipid portion of the GPI anchor with GPI-specific phospholipase C. Furthermore, the diversity of GPI anchored proteins that bind the toxin indicates that the binding occurs via the GPI

glycan without peptide requirements. Therefore, alpha toxin can be used as a lectin specific for the GPI anchor [51].

The primary mechanism that releases GPI anchored proteins from tumors into the circulatory system is not well understood. GPI anchored proteins can potentially be released from cells by proteolysis [53], GPI-specific phospholipase activities [54, 55], or by exosome vesicular transport from the cell membrane [53] (Fig. 1). The GPI glycan would remain attached to the GPI anchored proteins and bind with alpha toxin if the proteins were released by exosome or GPI-specific phospholipase cleavage (Fig. 1). However, GPI anchored proteins released by proteolysis would not bind alpha toxin due to lack of a GPI anchor. Our goal with the current study is to determine if alpha toxin can detect elevated levels of GPI anchored proteins present in plasma samples obtained for various human cancers.

PATIENTS AND METHODS

2.1 Patients

Blood from non-diseased and patients with ductal invasive breast carcinoma, ovarian cancer, kidney cancer, colon cancer, liver cancer, lung cancer, or brain cancer was collected pre-operatively and in accordance with approved institutional review board human subject guidelines at Georgia Health Sciences University (MCG) or the Ovarian Cancer Institute (Table 2). Plasma fractions were stored at -70°C until use.

2.2 Slot blot and alpha toxin detection

Plasma (5 ml) was mixed with 5 ml Laemelli buffer. Samples were heated to 95°C for 2 minutes to denature the proteins and cooled to room temperature before

application to the membrane. Nitrocellulose Protran BA85 membrane was immobilized and clamped securely using a Schleicher and Schuell Minifold I Slot Blot System. The samples were added to each slot before the vacuum was applied for 1 minute. Each well was washed using 1X PBS (200 ml) and again the vacuum was applied for 1 minute before removal of the membrane from the apparatus. The membrane was blocked in 5 % milk/1X TBST (Blotto Solution) overnight at 4°C. The blot was incubated with biotin labeled alpha toxin (2 mg/ml) expressed, purified, and labeled as described previously [51]. Bound toxin was detected using a 1:5,000 dilution of streptavidin-HRP (Vector Labs, Burlingame, CA) before washing and detection using Western Lightening Plus (Perkin Elmer). Slot blots were then stripped using 0.1M glycine pH 2.9 overnight, blocked again, and detected using anti-alpha 1 acid glycoprotein antibody (Sigma) to normalize for total protein content. Intensity of alpha toxin binding was determined using ImageJ analysis and was normalized to total protein band density.

2.3 Phospholipase C treatment and detection of CEA5 in LS174T cells.

Approximately 10 x 10^6 LS174T colon cancer cells were collected by gentle cell scraping. Cells were diluted with 200 ml 1X PBS supplemented with 5 mM calcium and magnesium chloride and evenly split into two fractions. One fraction received buffer only and one received 1.5 U/ml GPI-specific phospholipase C (Invitrogen) for 1 hour at 37°C. The cells were collected by centrifugation and the supernatants were collected for analysis. Biotin labeled alpha toxin ,2 mg/ml, was added and the samples were incubated at room temperature for 30 minutes. Streptavidin magnetic beads (Promega) were added for 30 additional minutes at room temperature. Beads were captured on a magnetic stand and washed 3X with 1X PBS before releasing the proteins with Laemmli

buffer. Proteins were separated on a 4-12 % Bis-Tris polyacrylamide gel (Invitrogen) and transferred to PVDF for detection of CEA5 (monoclonal antibody COL-1, Invitrogen) or gels were fixed and stained with silver.

RESULTS

In our previous study, we suppressed the expression of the *GPAA1* and *PIGT* subunits of the GPIT, enabling us to establish that alpha toxin binding required the addition of the GPI anchor to proteins [51]. Here, we show that alpha toxin can also bind to GPI anchored proteins that have been cleaved by GPI-specific phospholipase. As shown in Fig. 2A, LS174T colon cancer cells were incubated with or without GPI-specific phospholipase C. Biotin labeled alpha toxin was then added to the released proteins to capture GPI anchored proteins. Western blot detection of the GPI anchored protein carcinoembryonic antigen 5 (CEA5) indicate that CEA5 is present in the input supernatant samples from both untreated and PI-PLC treated cells. However, alpha toxin only captured CEA5 from the supernatant of PI-PLC treated cells indicating that the CEA5 released endogenously from LS174T cells does not contain the GPI anchor glycan and is released by proteolysis. Equivalent levels of proteins were present from the supernatants of both reactions (Fig. 2B); therefore, alpha toxin binding requires the presence of the GPI anchor glycan attached to GPI anchored proteins as found following PI-PLC release.

The levels of GPI anchored proteins present in plasma can be controlled by many factors, such as the levels of protein acceptors, the levels of GPIT subunits, the levels of endogenous GPI phospholipase activity, and the levels of protease activity. In addition

to these factors, solid tumors in different organs may sequester GPI anchored proteins into highly hydrophobic lipid raft membrane domains that are resistant to enzyme release. Based on our analysis of breast cancer tissue and plasma, increased levels of GPI anchored proteins were present in plasma from these patients. Breast cancers have frequent amplifications of chromosomal regions that contain the GPIT subunits. In Table 1, we list the subunits of the GPIT and the chromosomal location of each. In addition, we indicate human cancers that have chromosomal amplifications in the regions that contain GPIT subunits [56-66]. Furthermore, we cite existing published data indicating increased expression for GPIT subunits in certain human cancers [48-50, 67-69]. The information in Table 1 indicates that the GPI transamidase is elevated in many human cancers, therefore we hypothesize that increased levels of GPI anchored proteins are present in plasma from patients with these malignancies. To test this hypothesis, we have analyzed plasma collected from breast, ovarian, kidney, liver, lung, colon, and brain cancer patients by performing slot blot followed by alpha toxin detection (Table 2, Fig. 3). An example, shown in Fig. 3A, indicates that GPI anchored proteins could be detected in plasma from breast, ovarian, kidney, liver, lung, colon, and brain cancer with no detection in plasma from patients without malignant disease (Table 2, Fig. 3). Each slot contained equivalent levels of plasma proteins as evidenced by the alpha-1 acid glycoprotein levels (Fig. 3A, right). We analyzed 10 samples from each cancer type except ovarian (6 cases) along with 12 plasma samples from patients without malignant disease. The non-malignant samples were composed of age-matched female controls to the breast cancer cases analyzed (10) and 2 samples from presumed healthy individuals no sex and age information available. Densitometry

analysis of slot blots was performed, and the cumulative averaged alpha toxin signal intensities normalized to alpha-1 acid glycoprotein levels with SEM for each cancer are shown in Fig. 3B with value ranges shown for each sample in Table 2. These results indicate that GPI anchored proteins with a GPI anchor glycan attached were detected in the plasma from cancer patients at higher levels compared with plasma from patients without malignant disease. Variability exists in the levels of GPI anchored proteins detected by alpha toxin in different cancers and between different patients within certain cancer types. The cancers that show the highest variability are colon and brain cancer. Despite the patient variability, cumulative data indicate that GPI anchored proteins were detected in plasma at increased levels for all cancers analyzed (Fig. 3B). Therefore, proteomic studies to identify these GPI anchored proteins could lead to the discovery of novel biomarkers for cancer detection.

DISCUSSION

GPI anchored proteins are highly conserved and vital for viability [70-72]. However, the levels of GPI anchored proteins in normal cells are under tight control evidenced by the lower levels of GPIT mRNA and protein levels in normal tissues and cells [51]. GPIT levels are amplified in human cancers due to chromosomal amplifications acquired during malignant transformation. The impact of how the amplification of GPIT subunits can influence cancer progression is just beginning to be assessed. Breast cancer studies indicate that elevations of GPIT lead to increased levels of GPI anchored proteins [51] and increased levels of tumorigenicity [48]. Based on these findings, we sought to determine whether the levels of GPI anchored proteins in plasma for various

cancers correlates with previously described amplifications of GPIT levels in tissues for these cancers. Our results demonstrate that cancers with amplifications of certain GPIT subunits also have elevated alpha toxin binding to plasma from patients with these cancer types. We were surprised to discover that cancers with high GPIT mRNA and protein expression, such as breast and ovarian, do not have the highest levels of alpha toxin binding. For example, ovarian cancer has been shown to have the highest levels of expression for GPIT subunits, catalytic and non-catalytic [50]. Yet, our data reveals that while GPI anchored proteins are detected by alpha toxin at higher levels in ovarian cancer compared with control plasma, the levels are lower than other cancers. These results illustrate that other factors contribute to release of GPI anchored proteins into plasma in a form that can be detected by alpha toxin. These factors may include possible amplification of protease activity or release by an endogenous GPI phospholipase that may result in a modified GPI anchor glycan region that is not recognized as avidly by alpha toxin. We have obtained proteomic data that indicate high levels of GPI anchored protein expression detected by alpha toxin binding to membrane proteins extracted from ovarian cancer tumors (data not shown) and the surface of cells isolated from patient ascites [73]; therefore, future studies to detail the differences in the GPI anchor structures of these proteins from tissue, ascites, and serum may offer insight into why alpha toxin binding is lower for ovarian cancer patient plasma.

The functional significance of why GPI anchored proteins are released from the cell is not well understood. Studies from unicellular eukaryotic species have offered some insights into possible roles for released GPI anchored proteins [74]. Possible

explanations include: greater cell to cell communication, a method to control antigenic variability, and control of cell shape influencing growth and migration characteristics of cells. Therefore, it is not difficult to envision how tumor cells would gain a survival advantage by releasing GPI anchored proteins. We have discovered that cancers from the colon and brain have the highest levels of GPI anchored proteins detected in plasma. Plasma obtained from patients with glioblastoma brain tumors and lower grade colon adenocarcinomas show the highest levels of alpha toxin binding. Increased release of GPI anchored proteins from these cancers may reflect higher levels of cell to cell communication such as synaptic activity in the brain, or a greater need to diversify cell surface antigens as observed in colonic epithelial cells.

In conclusion, our data documenting elevated levels of GPI anchored proteins in the plasma from human cancers indicate that this type of glycoconjugate is a potentially useful biomarker for the monitoring and detection of human cancers. We have demonstrated that alpha toxin can be used as a GPI-specific lectin to detect GPI anchored proteins in human plasma. Therefore, the identification of specific GPI anchored proteins for these cancers can foster the development of novel cancer detection methods and possible therapeutic strategies utilizing alpha toxin in the future.

ACKNOWLEDGEMENTS

This work was supported by a Department of Defense Breast Cancer Concept Award to KLA (BC075534) in addition to funding from UO1CA128454 (MP), UO1CA168870 (KLA) and P41RR018502 (MP).

REFERENCES

1. Wu, G., Guo, Z., Chatterjee, A., Huang, X., Rubin, E., Wu, F., Mambo, E., Chang, X., Osada, M., Sook Kim, M., Moon, C., Califano, J. A., Ratovitski, E. A., Gollin, S. M., Sukumar, S., Sidransky, D., and Trink, B. Overexpression of glycosylphosphatidylinositol (GPI) transamidase subunits phosphatidylinositol glycan class T and/or GPI anchor attachment 1 induces tumorigenesis and contributes to invasion in human breast cancer. Cancer Res, *66:* 9829-9836, 2006.
2. Jiang, W. W., Zahurak, M., Zhou, Z. T., Park, H. L., Guo, Z. M., Wu, G. J., Sidransky, D., Trink, B., and Califano, J. A. Alterations of GPI transamidase subunits in head and neck squamous carcinoma. Mol Cancer, *6:* 74, 2007.
3. Nagpal, J. K., Dasgupta, S., Jadallah, S., Chae, Y. K., Ratovitski, E. A., Toubaji, A., Netto, G. J., Eagle, T., Nissan, A., Sidransky, D., and Trink, B. Profiling the expression pattern of GPI transamidase complex subunits in human cancer. Mod Pathol, *21:* 979-991, 2008.
4. Zhao, P., Nairn, A. V., Hester, S., Moremen, K. W., O'Regan, R. M., Oprea, G., Wells, L., Pierce, M., and Abbott, K. L. Proteomic identification of glycosylphosphatidylinositol anchor-dependent membrane proteins elevated in breast carcinoma. J Biol Chem., *287*: 25230-25240, 2012.
5. Eisenhaber, B., Bork, P., and Eisenhaber, F. Post-translational GPI lipid anchor modification of proteins in kingdoms of life: analysis of protein sequence data from complete genomes. Protein Eng, *14:* 17-25, 2001.
6. Fernandez-Messina, L., Ashiru, O., Boutet, P., Aguera-Gonzalez, S., Skepper, J. N., Reyburn, H. T., and Vales-Gomez, M. Differential mechanisms of shedding of the glycosylphosphatidylinositol (GPI)-anchored NKG2D ligands. J Biol Chem, *285:* 8543-8551.
7. Ikezawa, H. Bacterial PIPLCs-unique properties and usefulness in studies on GPI anchors. Cell Biol Int Rep, *15:* 1115-1131, 1991.
8. Lauc, G. and Heffer-Lauc, M. Shedding and uptake of gangliosides and glycosylphosphatidylinositol-anchored proteins. Biochim Biophys Acta, *1760:* 584-602, 2006.
9. Gorringe, K. L., George, J., Anglesio, M. S., Ramakrishna, M., Etemadmoghadam, D., Cowin, P., Sridhar, A., Williams, L. H., Boyle, S. E., Yanaihara, N., Okamoto, A., Urashima, M., Smyth, G. K., Campbell, I. G., and Bowtell, D. D. Copy number analysis identifies novel interactions between genomic loci in ovarian cancer. PLoS One, *5*.
10. Brown, K. M., Macgregor, S., Montgomery, G. W., Craig, D. W., Zhao, Z. Z., Iyadurai, K., Henders, A. K., Homer, N., Campbell, M. J., Stark, M., Thomas, S., Schmid, H., Holland, E. A., Gillanders, E. M., Duffy, D. L., Maskiell, J. A., Jetann, J., Ferguson, M., Stephan, D. A., Cust, A. E., Whiteman, D., Green, A., Olsson, H., Puig, S., Ghiorzo, P., Hansson, J., Demenais, F., Goldstein, A. M., Gruis, N. A., Elder, D. E., Bishop, J. N., Kefford, R. F., Giles, G. G., Armstrong, B. K., Aitken, J. F., Hopper, J. L., Martin, N. G., Trent, J. M., Mann, G. J., and Hayward, N. K. Common sequence variants on 20q11.22 confer melanoma susceptibility. Nat Genet, *40:* 838-840, 2008.
11. Ishihara, T., Tsuda, H., Hotta, A., Kozaki, K., Yoshida, A., Noh, J. Y., Ito, K., Imoto, I., and Inazawa, J. ITCH is a putative target for a novel 20q11.22 amplification

detected in anaplastic thyroid carcinoma cells by array-based comparative genomic hybridization. Cancer Sci, *99:* 1940-1949, 2008.
12. Fujita, Y., Sakakura, C., Shimomura, K., Nakanishi, M., Yasuoka, R., Aragane, H., Hagiwara, A., Abe, T., Inazawa, J., and Yamagishi, H. Chromosome arm 20q gains and other genomic alterations in esophageal squamous cell carcinoma, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization. Hepatogastroenterology, *50:* 1857-1863, 2003.
13. Hidaka, S., Yasutake, T., Kondo, M., Takeshita, H., Yano, H., Haseba, M., Tsuji, T., Sawai, T., Nakagoe, T., and Tagawa, Y. Frequent gains of 20q and losses of 18q are associated with lymph node metastasis in intestinal-type gastric cancer. Anticancer Res, *23:* 3353-3357, 2003.
14. Haiman, C. A., Patterson, N., Freedman, M. L., Myers, S. R., Pike, M. C., Waliszewska, A., Neubauer, J., Tandon, A., Schirmer, C., McDonald, G. J., Greenway, S. C., Stram, D. O., Le Marchand, L., Kolonel, L. N., Frasco, M., Wong, D., Pooler, L. C., Ardlie, K., Oakley-Girvan, I., Whittemore, A. S., Cooney, K. A., John, E. M., Ingles, S. A., Altshuler, D., Henderson, B. E., and Reich, D. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet, *39:* 638-644, 2007.
15. Yeager, M., Orr, N., Hayes, R. B., Jacobs, K. B., Kraft, P., Wacholder, S., Minichiello, M. J., Fearnhead, P., Yu, K., Chatterjee, N., Wang, Z., Welch, R., Staats, B. J., Calle, E. E., Feigelson, H. S., Thun, M. J., Rodriguez, C., Albanes, D., Virtamo, J., Weinstein, S., Schumacher, F. R., Giovannucci, E., Willett, W. C., Cancel-Tassin, G., Cussenot, O., Valeri, A., Andriole, G. L., Gelmann, E. P., Tucker, M., Gerhard, D. S., Fraumeni, J. F., Jr., Hoover, R., Hunter, D. J., Chanock, S. J., and Thomas, G. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet, *39:* 645-649, 2007.
16. Daley, D., Lewis, S., Platzer, P., MacMillen, M., Willis, J., Elston, R. C., Markowitz, S. D., and Wiesner, G. L. Identification of susceptibility genes for cancer in a genome-wide scan: results from the colon neoplasia sibling study. Am J Hum Genet, *82:* 723-736, 2008.
17. Hashimoto, K., Mori, N., Tamesa, T., Okada, T., Kawauchi, S., Oga, A., Furuya, T., Tangoku, A., Oka, M., and Sasaki, K. Analysis of DNA copy number aberrations in hepatitis C virus-associated hepatocellular carcinomas by conventional CGH and array CGH. Mod Pathol, *17:* 617-622, 2004.
18. Fabris, S., Todoerti, K., Mosca, L., Agnelli, L., Intini, D., Lionetti, M., Guerneri, S., Lambertenghi-Deliliers, G., Bertoni, F., and Neri, A. Molecular and transcriptional characterization of the novel 17p11.2-p12 amplicon in multiple myeloma. Genes Chromosomes Cancer, *46:* 1109-1118, 2007.
19. Weiss, M. M., Kuipers, E. J., Postma, C., Snijders, A. M., Pinkel, D., Meuwissen, S. G., Albertson, D., and Meijer, G. A. Genomic alterations in primary gastric adenocarcinomas correlate with clinicopathological characteristics and survival. Cell Oncol, *26:* 307-317, 2004.
20. Guo, Z., Linn, J. F., Wu, G., Anzick, S. L., Eisenberger, C. F., Halachmi, S., Cohen, Y., Fomenkov, A., Hoque, M. O., Okami, K., Steiner, G., Engles, J. M., Osada, M., Moon, C., Ratovitski, E., Trent, J. M., Meltzer, P. S., Westra, W. H., Kiemeney, L. A., Schoenberg, M. P., Sidransky, D., and Trink, B. CDC91L1 (PIG-U) is a newly discovered oncogene in human bladder cancer. Nat Med, *10:* 374-381, 2004.

21. Shen, Y. J., Ye, D. W., Yao, X. D., Trink, B., Zhou, X. Y., Zhang, S. L., Dai, B., Zhang, H. L., Zhu, Y., Guo, Z., Wu, G., and Nagpal, J. Overexpression of CDC91L1 (PIG-U) in bladder urothelial cell carcinoma: correlation with clinical variables and prognostic significance. BJU Int, *101:* 113-119, 2008.
22. Ho, J. C., Cheung, S. T., Patil, M., Chen, X., and Fan, S. T. Increased expression of glycosyl-phosphatidylinositol anchor attachment protein 1 (GPAA1) is associated with gene amplification in hepatocellular carcinoma. Int J Cancer, *119:* 1330-1337, 2006.
23. Leidich, S. D., Drapp, D. A., and Orlean, P. A conditionally lethal yeast mutant blocked at the first step in glycosyl phosphatidylinositol anchor synthesis. J Biol Chem, *269:* 10193-10196, 1994.
24. Kawagoe, K., Kitamura, D., Okabe, M., Taniuchi, I., Ikawa, M., Watanabe, T., Kinoshita, T., and Takeda, J. Glycosylphosphatidylinositol-anchor-deficient mice: implications for clonal dominance of mutant cells in paroxysmal nocturnal hemoglobinuria. Blood, *87:* 3600-3606, 1996.
25. Ferguson, M. A. The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. J Cell Sci, *112 (Pt 17):* 2799-2809, 1999.
26. Dangaj, D., Abbott, K. L., Mookerjee, A., Zhao, A., Kirby, P. S., Sandaltzopoulos, R., Powell, D. J., Jr., Lamaziere, A., Siegel, D. L., Wolf, C., and Scholler, N. Mannose receptor (MR) engagement by mesothelin GPI anchor polarizes tumor-associated macrophages and is blocked by anti-MR human recombinant antibody. PLoS One, *6:* e28386.
27. Simon, M. C. and Kusch, J. Communicative functions of GPI-anchored surface proteins in unicellular eukaryotes. Crit Rev Microbiol.

GPIT Subunit	Chromosome Location	Cancers with Amplifications of Chromosome Region	Cancers with Demonstrated Overexpression of GPIT Subunit	References
PIGU	20q11.22	bladder, breast, melanoma, ovarian, thyroid	bladder, breast, colon, head and neck, kidney, lung, ovarian, skin, thyroid, uterus	Hyman E et al. 2002, Gorringe, kl et al. 2010, Guo Z et al., 2004, Brown et al., 2008, Yi-Jun Shen et al., 2007, Jiang et al., 2007, and Ishihara et al., 2008
PIGT	20q12-q13.12	breast, esophageal, gastric, ovarian	breast, colon, head and neck, ovarian, skin, thyroid, uterus	Fujita et al. 2003, Hidaka S et al. 2003, Sonoda G et al. 1997, Wu et al. 2006, Nagpal et al. 2008
GPAA1	8q24.3	breast, colon, esophageal, hepatocellular, lung, prostate, uterine	breast, bladder, head and neck, hepatocellular, kidney, lung, ovarian, prostate, uterine	Hyman E et al. 2002, Mark HF et al. 1999, Nagpal et al. 2008, Wu et al. 2006, Ho JC et al. 2006, Jiang et al. 2007
PIGK(GP18)	1p31.1	colon, hepatocellular	bladder, breast, colon, liver, ovarian, prostate	Hashimoto et al. 2004, Daley D et al. 2008, Nagpal et al. 2008
PIGS	17p13.2	gastric, multiple myeloma	bladder, breast, lung, ovarian, uterine	Fabris S et al. 2007, Weiss MM et al. 2004, Nagpal et al. 2008

Table Supplement 1. GPI TRANSAMIDASE SUBUNITS AMPLIFIED IN HUMAN CANCERS

Cancer Type	*Histology*	*Grade*	*Number of Cases*	*Alpha Toxin Binding Value Range*
Breast	Invasive Ductal	I	2	0.8-2.4
		II	5	0.8-2.4
		III	3	0.8-2.4
Ovarian	Serous	III	3	0.2-1.0
	Endometrioid	III	3	0.2-1.0
Kidney	Papillary Adenocarcinoma	II	2	1.4-2.2
	Renal Cell Carcinoma	II	8	2.9-6.0
Liver	Adenocarcinoma	II	5	1.0-2.3
	Cholangiocarcinoma	III	1	2.3
	Carcinoid	unknown	3	3.5-7.0
	Sarcoma	unknown	1	1.9
Lung	Adenocarcinoma	I	1	2.2
		II	1	1.5
		unknown	2	2.1-2.5
	Squamous Cell	II	3	3.1-6.0
	Non Small Cell	III	1	2
	Malignant Mesothelioma	unknown	1	1.5
Colon	Adenocarcinoma	I	1	12
		II	8	2.1-4.0
	Adenocarcinoma in tubulovillous adenoma	II	1	9
Brain	Meningioma	II	3	4.1-6.1
	Glioblastoma	IV	2	8.1-9.0
	Oligosarcoma	unknown	1	3.2
	Gliosarcoma	IV	1	2.1
	Adenocarcinoma	unknown	1	2.1
	Glioma	II	1	1.2
Non-Malignant			12	0.1-0.2

Table Supplement 2. ALPHA TOXIN BINDING SUMMARY FOR PLASMA FROM PATIENTS WITH HUMAN CANCER AND CONTROLS

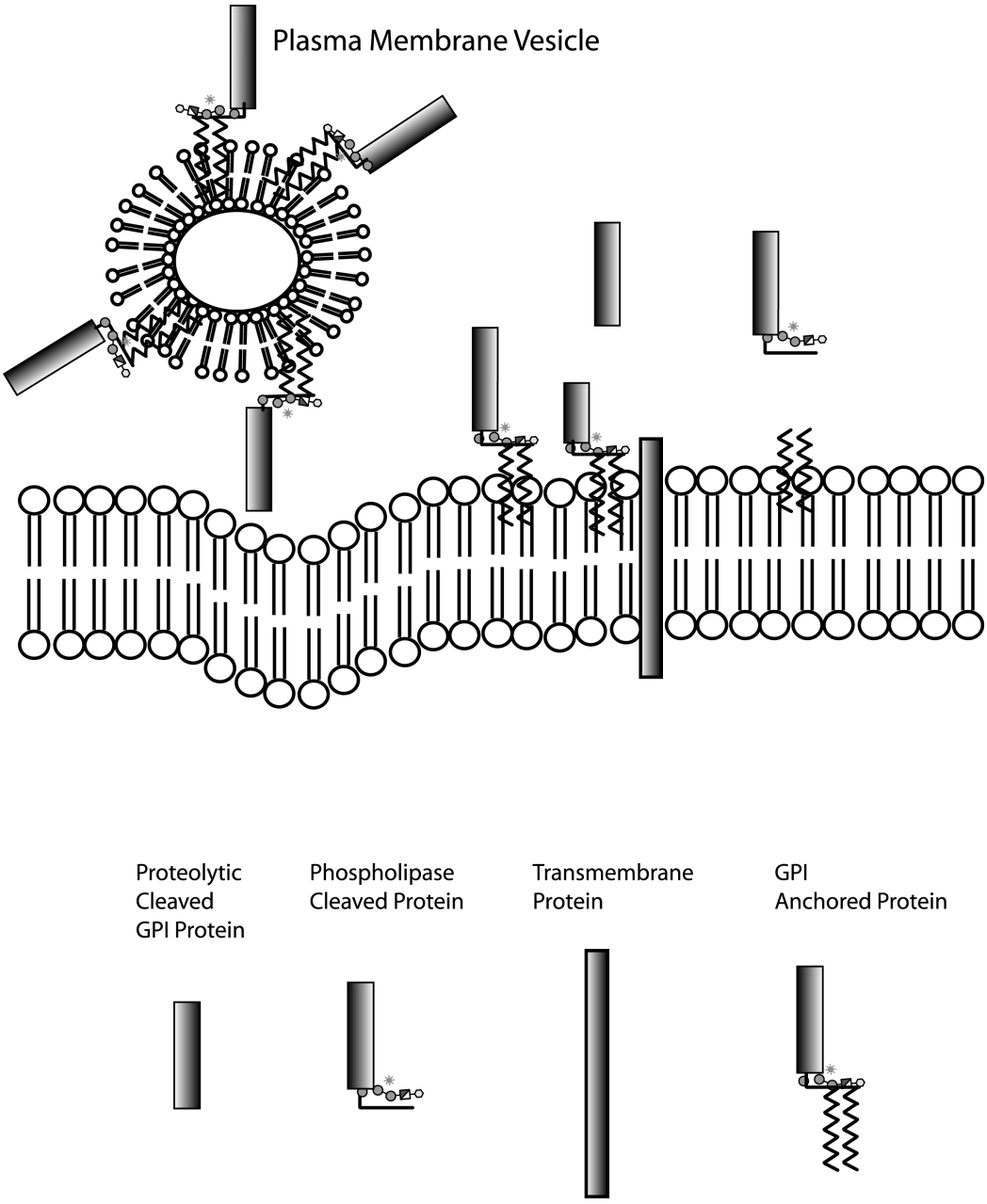


Figure Supplement 1. DIAGRAM ILLUSTRATING VARIOUS MECHANISMS THAT CAN RESULT IN THE RELEASE OF GPI ANCHORED PROTEINS FROM THE CELL SURFACE. GPI anchored proteins released from membrane-derived vesicles or via GPI-specific phospholipase enzymes would be expected to have an intact GPI anchor (lipid and glycan) or partial GPI anchor (glycan). GPI anchored proteins released by proteolysis will not have the GPI anchor.

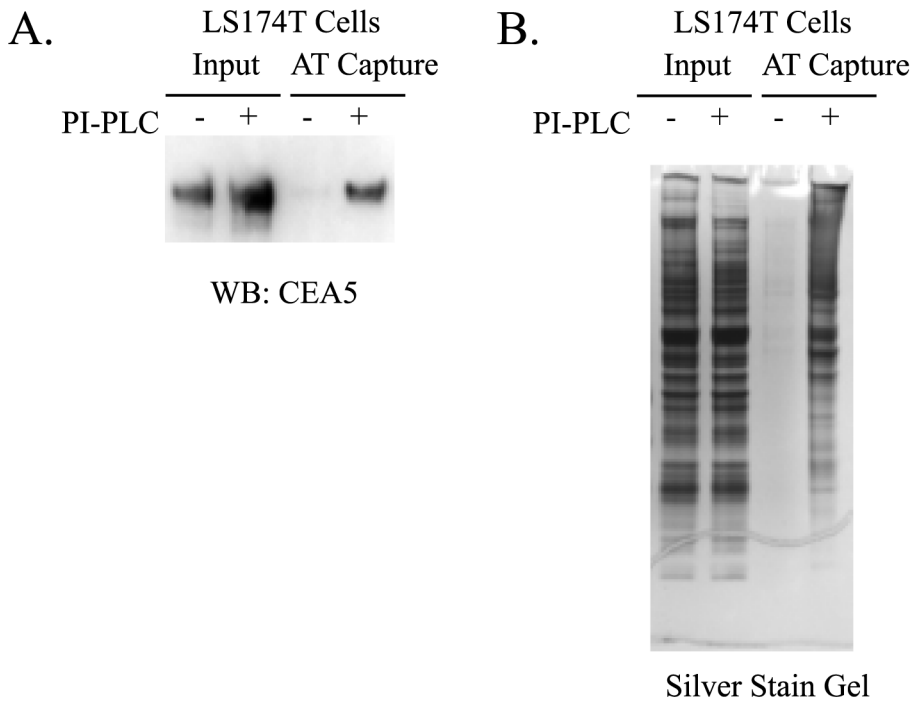
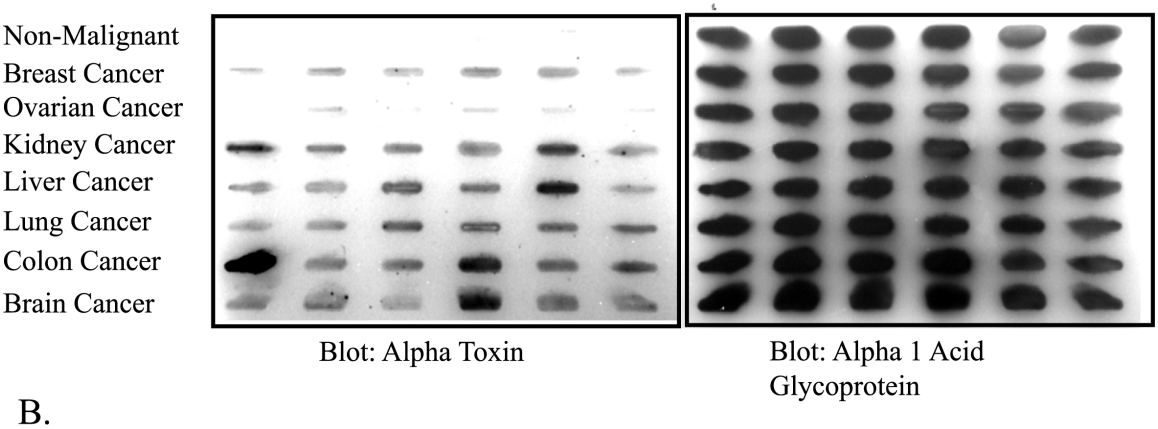


Figure Supplement 2. ALPHA TOXIN CAPTURE OF CEA5 REQUIRES THE PRESENCE OF THE GPI ANCHOR. (A) LS174T human colon adenocarcinoma cells were divided into equal fractions, one fraction received buffer only (-) and one fraction received GPI-specific PI-PLC (1.5 units/ml) (+). Following incubation for one hour at 37°C, the supernatant was subjected to alpha toxin capture followed by Western blot detection of carcinoembryonic antigen 5 (CEA5). Input equals 10 % of the supernatant used for alpha toxin capture. (B) Ten percent protein inputs and alpha toxin captured proteins were separated on a 4-12 % polyacrylamide gel and stained with silver.

A.



B.

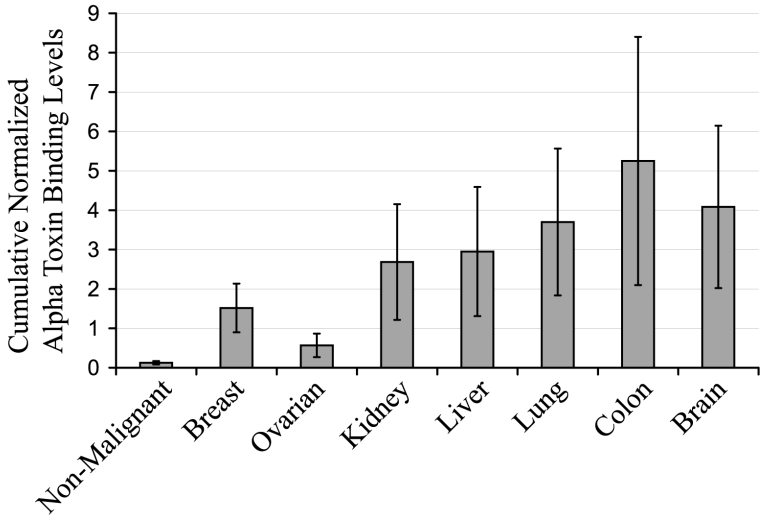


Figure Supplement 3. ALPHA TOXIN REACTS WITH PLASMA FROM HUMAN CANCERS AND NOT WITH CONTROL (NON-MALIGNANT) PLASMA. (A) Representative slot blot analysis of human plasma (5 mL) from patients with the indicated cancers detected with biotin labeled alpha toxin (2 mg/ml) followed by streptavidin conjugated peroxidase and chemiluminescent substrate. (B) Cumulative results from slot blot analysis. Alpha toxin binding levels were normalized to alpha-1-acid glycoprotein levels for each case. Error bars indicate the ±SEM for normalized densities obtained from all serum samples analyzed for each cancer (listed in Table 2).