Universidade de Aveiro
2012

Departamento de Química

ANDREIA FILIPA FERREIRA DE ALMEIDA

Deteção de O-glicanos sialilados associados ao tumor por MALDI-TOF/TOF

Universidade de Aveiro
2012

Departamento de Química

ANDREIA FILIPA FERREIRA DE ALMEIDA

Deteção de O-glicanos sialilados associados ao tumor por MALDI-TOF/TOF

Detection of tumor-associated sialylated O-glycans by MALDI-TOF/TOF

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica – clínica, realizada sob a orientação científica do professor Doutor Celso Albuquerque Reis, professor no Instituto de Patologia e Imunologia da Universidade do Porto (IPATIMUP) e do Doutor José Alexandre Ribeiro de Castro Ferreira, Investigador de pós-doutoramento do Centro de Investigação do Instituto Português de Oncologia do Porto (IPO-Porto) e do Departamento de Química da Universidade de Aveiro.

This work is dedicated to all my family, friends and supervisors

The judges

Presidente

Professora Doutora Maria do Rosário Gonçalves Reis Marques Domingues
Professora auxiliar do departamento de Química da Universidade de Aveiro

Doutor Hugo Alexandre de Carvalho Pinheiro Osório
Investigador auxiliar no Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP)

Doutor José Alexandre Ribeiro de Castro Ferreira
Investigador de Pós-Doutoramento do Departamento de Química da Universidade de Aveiro e do Centro de Investigação do Instituto Português de Oncologia do Porto (IPO-Porto).

Professor Doutor Celso Albuquerque Reis
Professor no Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP).

Palavras-chave

O-glicosilação, antigénios associados ao antigénio Thomsen-Friedenreich, cancro da bechiga, lesões precursoras do cancro gástrico, MALDI, ácidos siálicos, permetilação

Resumo

Uma das alterações fenótipicas mais comuns nos tumores são modificações no padrão de O-glicosilação na superfície da célula e nas glicoproteínas secretadas. Como consequência têm implicações nas suas funções biológicas. Em particular, tem sido descrito que algumas células tumorais sobreexpressam ou expressam *de novo* antigénios associados ao antigénio Thomsen-Friedenreich (TF), ou seja, sialil-Tn, sialil-T e disialyl-T. Estes epítopos resultam da paragem prematura do processo de O-glicosilação em proteínas e têm um impacto direto sobre a biologia do tumor. Assim sendo, a identificação destas modificações pós-translacionais anormais em proteínas é essencial para determinar as relações estrutura-função e descobrir novos alvos terapêuticos. Além disso, as proteínas que transportam estas alterações podem ser secretadas na corrente sanguínea, urina, e em outros fluidos corporais e, portanto, são explorados como biomarcadores em testes não invasivos. Atualmente a detecção de antigénios associados ao antigénio TF é baseado em métodos imunohistoquímicos em que, embora úteis para uma investigação de rotina, não conseguem descrever totalmente o padrão de glicosilação de uma dada proteína. Sendo assim, neste trabalho apresentamos uma abordagem analítica para determinar estes glicanos em quantidades minimas de glicoproteínas (*pico*mole) isoladas a partir de géis SDS-PAGE. Resumidamente, as glicoproteínas são de-O-glicosiladas no gel por beta-eliminação redutiva, permetiladas e analisadas por nanoLC-MALDI-TOF/TOF. De seguida, os dados provenientes são sujeitos a uma seleção melhorada dos sinais analíticos relevantes, utilizando ferramentas de bioinformática. Esta abordagem foi, em seguida, aplicada com sucesso na validação do western blotting quanto à expressão de sialil-Tn numa glicoproteína isolada a partir da urina de ratos com tumores na bexiga induzidos quimicamente e no plasminogénio isolado a partir do soro de pacientes com lesões precursoras do cancro gástrico.

keywords

O-glycosylation, Thomsen-Friedenreich-related antigens, bladder tumors, gastric cancer precursor lesions, MALDI, sialic acids, permethylation

Abstract

A common phenotypic change in tumors comprises alterations in the O-glycosylation of cell-surface and secreted glycoproteins with implications in their biological functions. In particular, it has been described that some tumor cells overexpress or *de novo* express Thomsen-Friedenreich (TF)-related antigens, namely sialyl-Tn, sialyl-T and disialyl-T. These epitopes result from a premature stop in protein O-glycosylation and have direct impact on tumor biology. As a result, the identification of these abnormal post-translational modifications of proteins is essential to determine structure-function relationships and designs novel therapeutics. Moreover, the proteins carrying these alterations can ultimately be shed into the blood stream, urine and other body fluids and thus be explored as biomarkers in non invasive tests. Currently the detection of TF-related antigens relies on immuno-based methods that, even though useful in a routine basis, often fail to fully highlight the glycosylation pattern of a given protein. Herein, we have systematized a target-driven analytical approach to determine these glycans in minute amounts of glycoproteins (*pico*mole) isolated from SDS-PAGE gels. Briefly, the glycoproteins are to be de-O-glycosylated in-gel by reductive beta-elimination, permethylated and analyzed by nanoLC-MALDI-TOF/TOF with enhanced selection of the relevant analytical signals using bioinformatics tools. This approach was then successfully applied to validate western blotting assignments regarding the expression of sialyl-Tn in a glycoprotein isolated from the urine of rats with chemically-induced bladder tumors and in plasminogen isolated from the serum of patients with gastric cancer precursor lesions.

Contents

Chapter I. State of the art **1**

1. Introduction 3
2. Structure and biosynthesis of O-GalNAc glycans in mammalian cells 5
3. Alterations at O-glycosylation level in tumors 9
4. Analytical strategies to analyze sialylated T-related O-glycans by Mass Spectrometry 10
   - 4.1. Methods for isolation of O-GalNAc glycoproteins 11
   - 4.2. Analysis of sialylated glycans by permethylation 12
   - 4.3. Collision Induced Dissociation of oligosaccharides 14
   - 4.4. CID of permethylated oligosaccharides 16
5. References 17

Chapter II. Aims and Scopes 26

Analytical approaches 27

Chapter III. Detection of sialylated cancer-associated Thomsen-Friedenreich - related antigens by in gel deglycosylation and nanoLC-MALDI-TOF/TOF 28

Abstract 29

Introduction 30

Experimental 32

Results and Discussion 38

Conclusion 50

References 51

Chapter IV. Detection of sialyl-Tn in plasminogen from serum of patients with gastric cancer precursor lesions by in gel deglycosylation and nanoLC-MALDI TOF/TOF **56**

Abstract 57

Introduction 58

Experimental methods 59

Results and Discussion 61

Conclusion 63

References 64

Chapter V. Discussion and Conclusion **66**

References 68

Figure captions

Chapter I. State of the art..... **1**

Figure 1. The biosynthetic pathway of O-GalNAc glycosylation 8

Figure 2. Reduction of oligosaccharides with sodium borohydride and further permethylation..... 14

Figure 3. The nomenclature for oligosaccharide fragmentation..... 15

Chapter III. Detection of sialylated cancer-associated Thomsen-Friedenreich - related antigens by in gel deglycosylation and nanoLC-MALDI-TOF/TOF 28

Figure 1. a) Electrophoretic profile of fetuin in 4-12% SDS-PAGE highlighting the amount of glycoprotein used for in-gel de-O-glycosylation, b) Positive MALDI-MS spectra of permethylated O-glycans released from 4 *pico*mol of fetuin with an expansion highlighting the ions assigned in Table 1, and b) expansion of the zones of the spectrum were sialyl-Tn, sialyl-T (s-3-T and s-6-T) and disialyl-T ions are expected 40

Figure 2. Plots of the sparse matrix (Fraction Number *vs m/z vs* nano-HPLC-MALDI-MS signal) for the peaks found within 873.5±0.5 Da (a and c) and 895.5±0.5Da (b and d), which includes the monoisotopic mass of protonated and sodiated forms of sialyl-T, respectively. Plots a) and b) refer to O-glycans resulting from 4 *pico*mol of native fetuin, whereas c) and d) refer to fetuin subjected to treatment with an α-neuraminidase. A chromatographic envelop is visually detected in figures a) and c) both due to the existence of a cluster of high-intensity signals, but also by the lesser dispersion of *m/z* over the envelope region..... 43

Figure 3. Figure 1. Nano-HPLC-MALDI-MS fractions exhibiting distinguishable signals for a) sialyl-Tn, b) sialyl-T (s-3-T and s-6-T), and c) disialyl-T $[M+H]^+$ and $[M+Na]^+$ ions.. 44

Figure 4. MALDI-TOF/TOF CID spectra of $[M+H]^+$ permethylated a) sialyl-Tn, b) sialyl-6-T, and c) disialyl-T exhibiting characteristic fragmentations according to the nomenclature introduced by Domon and Costello (1988). "*" signals resulting from the combination of cross-ring and glycosidic bound fragmentations occurring at both the reducing and non-reducing ends.. 46

Figure 5. MALDI-TOF/TOF CID spectra of $[M+Na]^+$ permethylated a) sialyl-Tn, b) sialyl-T, and c) disialyl-T exhibiting characteristic fragmentations according to the nomenclature introduced by Domon and Costello (1988). "*" signals resulting from the combination of cross-ring and glycosidic bound fragmentations occurring at both the reducing and non-reducing ends. 47

Figure 6. a) Rat urothelium with BBN-induced tumor showing intense immunostaining for sialyl-Tn; b) SDS-PAGE gel section from urine proteins showing a tumor specific band at 25 kDa (I) immunoreactive with anti-sialyl Tn monoclonal antibody TKH2 (II); c) Plots of the sparse matrix (Fraction Number *vs m/z vs* nano-HPLC-MALDI-MS signal) for the peaks found within 669.4±0.2 Da, which includes the monoisotopic mass of protonated form of sialyl-Tn.. 50

Chapter IV. Detection of sialyl-Tn in plasminogen from serum of patients with gastric cancer precursor lesions by in gel deglycosylation and nanoLC-MALDI TOF/TOF 56

Figure 1. a) SDS-PAGE gel showing plasminogen from serum of healthy donors (Ctrl) and patients diagnosed with gastritis (Gast), complete (MIC) and incomplete (MII) intestinal metaplasia isolated by lysine-sepharose affinity chromatography and b) nano-HPLC-MALDI-MS fraction of permethylated sialyl-Tn from plasminogen of patients with incomplete metaplasia intestinal exhibiting $[M+Na]^+$ and c) MALDI-TOF/TOF spectra of $[M+Na]^+$ permethylated sialyl-Tn with unique fragmentations according to the nomenclature of Domon and Costello (1988). "*****" signals resulting from the combination of cross-ring and glycosidic bound fragmentations occurring at both the reducing and non-reducing ends..... 62

Tables

Chapter I. State of the art..... **1**

Table 1. Structures of O-GalNAc cores. 7

Chapter III. Detection of sialylated cancer-associated Thomsen-Friedenreich - related antigens by in gel deglycosylation and nanoLC-MALDI-TOF/TOF 28

Table 1. Permethylated glycans generated by chemical de-O-glycosylation of 4 *pico*mol from fetuin showing MALDI-MS ions with a signal-to-noise ratio superior to 10. 41

Chapter I. State of art

Abbreviations

β4/3GalTs - β1,-4/3 Galactosyltranferases
β3/4GnTs - β1-3/4 *N*-acetylglucosaminyltransferases
C1GalT - Core 1 β1-3 Galactosyltransferase or T synthase
C2GnT - Core 2 β1-6 *N*-acetylglucosaminyltransferase
C3GnT - Core 3 β1-3 *N*-acetylglucosaminyltransferase
CID- Colision induced dissocation
CMP-Neu5Ac - Cytosine monophosphate *N*-acetyl neuraminic acid
Cosmc - Core 1 β1-3galactosyltranferase-specific molecular chaperone
ER - Endoplasmic Reticulum
Fuc – Fucose
Gal - Galactose
GalNAc – *N*-acetyl-galactosamine
GlcNAc - *N*-acetyl-glucosamine
HexNAc - *N*-acetyl-hexosamine
MALDI - Matrix-assisted laser desorption/ionization
MS- Mass spectrometry
Neu5Ac – *N*-acetyl-neuraminic acid
ppGalNAc-Ts - polypeptide *N*-acetylgalactosaminyltransferases
Pro – Poline
PTM - post-translational modification
Ser - Serine
ST3Gal - α 2-3 Sialyltransferases galactosamine
ST6GalNAc - α 2-6 Sialyltransferases *N*-acetylgalactosamine
Thr – Threonine
TOF - Time of flight
UDP-Gal - Uridine diphosphate - galactosamine
UDP-GalNAc - Uridine diphosphate - *N*-acetylgalactosamine
UDP-GlcNAc - Uridine diphosphate - *N*-acetylglucosamine

1. Introduction

Glycosylation is one of the most ubiquitous and complex post-translational modification (PTM) and is estimated to occur on more than half of the proteins encoded in eukaryotic human genomes [1]. The glycosylation of mammalian proteins is achieved by the action of glycosyltransferases, that catalyze the transfer of a sugar from a donor molecule to the a potential site of polypeptide or carbohydrate [2]. During malignant transformation of the cell this "glycosylation machinery" is significantly compromised. Such changes have been explained by either a disorganization of secretory pathway organelles (Endoplasmic reticulum and Golgi apparatus) and/or altered glycosyltransferase expression [3]. A deregulation in fine tuning of this "glycosylation machinery" has direct impact in the cell biological behavior, namely its adhesion, signaling properties, and interaction with the immune system [4]. Among the most common alterations in O-gly*c*osylation is the expression of short oligosaccharide chains, the Thomsen-Friedenreich(TF)-related antigens. This low molecular weight glycans result from a premature stop in biosynthesis of O-glycans, leading to the formation of Tn (GalNAcα-O-Ser/Thr) and T (Galβ1-3GalNAcα-O-Ser/Thr) antigens. Tn an T antigens are precursors for many complex structures, that ultimately can be terminated with Lewis and blood group determinants [5]. In the presence of specific sialyltransferases they may form their sialylated-counterparts, sialyl Tn (Neu5Acα2-3GalNAcα-O-Ser/Thr), sialyl-T (Neu5Acα2-3Galβ1-3GalNAcα-O-Ser/Thr), sialyl-6-T (Galβ1–3(NeuAcα2-6)GalNAcα-O-Ser/Thr) and disialyl-T (NeuAcα2-3Galβ1-3[NeuAcα2-6]GalNAcα-O-Ser/Thr). T-related antigens are also classically referred to as simple mucin type carbohydrates, since they are abundantly expressed in mucins due to the high density of O-glycosylation sites [3].

Sialylated TF-related antigens are generally expressed in low levels or not expressed at all in healthy tissues; yet they are frequently (over)expressed in several human malignancies [6]. The sialyl-Tn is the most studied epitope due its overexpression and *de novo* biosynthesis in several cancers, including gastric tumors [7]. Thereby it may be considered as a hallmark of cancer cell transformation, correlating with poor prognosis, poor outcome, aggressiveness and metastization [2, 3, 5]. As a result it has been targeted for immunotherapy and explored both as a biomarker and in cancer vaccines [8]. The

Sialyl-T and disialyl-T are commonly expressed in the healthy human epithelium, but still overexpressed in some malignancies, such as bladder tumors [9].

The role of the sialylated TF-related antigens, in particular sialyl-Tn, in human health and disease has been a hot topic over the last twenty to thirty years and regarded of great interest for carbohydrate-based vaccines, drug development and diagnosis [10]. The complexity of these glycans and its association with proteins, makes of structural characterization a key part of in the context of research. Nowadays mass spectrometry (MS) has proven to be the core technology behind proteomics, and it stands to play a similar role in the study of the functional implications of carbohydrate expression [11, 12]. Mass spectrometry may provide insights on the purity, molecular mass, monosaccharide composition, sequence of sugar residues, presence and position of branches, inter-glycosidic linkages and isomeric molecules [13]. In particular, matrix-assisted laser desorption/ionization (MALDI), is a soft ionization technique with high sensitivity, resolution and suitable for small amounts of sample. Even though widely used in proteomic research, MALDI ionization is known to transfers high energy to labile compounds, such as sialic acids, promoting their fragmentation [14, 15]. This has been mostly overcome by derivatization approaches such as permethylation, that provides an increase ionization response allowing their detection in the positive mode [12, 16-18]. Moreover, it provides highly informative product ion spectra that may allow a full assignment of the primary structure, including the distinction of isomers [19]. However, the high pH used in permethylation promotes a significant degree of oxidative degradation and peeling reactions [20, 21]. Undesirable by-products are also promoted by the high pH, reductive environment and the temperature necessary for the de-O-glycosylation of proteins in the absence of specific O-glycosidases. This synergism of factors is responsible by the introduction of significant background noise and, in some extent, the degradation of the glycans under analysis [22]. As such, the analysis by MALDI of sialylated TF-related antigens, having as starting material low *pico*mole to *femto*mole of a given glycoprotein, is currently regarded as major analytical challenge.

2. Structure and biosynthesis of O-GalNAc glycans in mammalian cells

The biosynthesis of the carbohydrate structures in glycoproteins relies on a number of competitive and very concerted processes involving several glycosyltransferases. Mucin (GalNAc)-type O-glycosylation (here called O-glycosylation) is one of the most common type of glycosylation found in glycoproteins and consists of a glycan *O*-linked to a serine (Ser) or a threonine (Thr) residue. The association of this type of glycosylation with mucins results from the high density of Ser/Thr/Pro domains encountered in this heavily glycosylated class of protein [23].

The first step of O-glycosylation involves the transfer of *N*-acetyl-galactosamine from a sugar donor UDP-GalNAc on to serine or threonine residue of a given polypeptide chain. This step is controlled by the activity of polypeptide *N*-acetyl-galactosaminyltransferases (ppGalNAc-Ts) and results in the formation of the Tn antigen (GalNAcα-O-Ser/Thr; Figure 1 and Table 1). Until now at least twenty one ppGalNAc-Ts isoforms have been identified in mammals [24, 25]. These class of enzymes determine both O-GalNAc sites and density of glycosylation [26] and are characterized by different tissue expression [27], kinetic properties and acceptor substrate specificities [28, 29]. All ppGalNAc-Ts bind UDP-GalNAc, but ppGalNAc-Ts are known to have high overlapping substrate specificity and many substrates are modified by several ppGalNAc-Ts [30].

The addition of the first GalNAc sugar into a region rich in Ser/Thr of the protein, changes the way other ppGalNAc-T isoforms act [31]. As such, the competition between ppGalNAc-Ts can cause different patterns of O-glycosylation. Therefore, the multiplicity of ppGalNAc-T isoforms fine-tunes the pattern and density of GalNAc modification on proteins, that are highly enriched in acceptor sites. Thus, co-expression in the same cell of ppGalNAcTs with complementary, partly overlapping acceptor substrate specificities probably ensures efficient O-GalNAc glycosylation [32]. After the first step of glycosylation, the Tn antigen can be prematurely terminated by sialylation when increased levels of ST6GalNAc-I/II are present. These enzymes are responsible by the transfer of a *N*-acetyl-neuraminic acid (Neu5Ac) from cytosine monophosphate *N*-acetyl-neuraminic acid (CMP-Neu5Ac) to the O-6 position of the GalNAc residue of a Tn antigen, thus originating the sialyl-Tn antigen (Neu5Acα2-6GalNAcα-O-Ser/Thr; Figure 1) [33]. The sialylation of the Tn antigen can be performed by two different isoforms of ST6GalNAc,

which are ST6GalNAc-I and ST6GalNAc-II. Still it has been demonstrated that, when the Tn and T antigens are present *in vitro* essays, ST6GalNAc-I acts preferentially on Tn antigen [34].

Despite these observations, in most mammalian tissues, the Tn antigen is a substrate for core 1 β1-3 galactosyltransferases (C1GalT) or T synthase. The formation of the active core 1 β1-3 galactosyltransferase *in vivo* is dependent on core 1 β1-3 galactosyltransferase-specific molecular chaperone (Cosmc) [35]. Cosmc is an Endoplasmic Reticulum (ER) protein that appears to bind specifically to T synthase and ensures its full activity in the Golgi Apparatus [36]. C1GalT catalyzes the transfer of a Gal from uridine diphosphate-galactose (UDP-Gal) to the O-3 position of GalNAc residue, generating the core 1 or T antigen (Galβ1-3GalNAcα-O-Ser/Thr; Figure 1 and Table 1). The T antigen can also be prematurely ended by sialylation. Unlike for sialyl-Tn, the sialylation of the T antigen is mostly mediated by α2-3 sialyltransferases, namely ST3Gal-I or ST3Gal-II [37]. The pointed out enzymes are responsible by the addition of a Neu5Ac from the sugar donor CMP-Neu5Ac to the O-3 position of the Gal residue of core 1 originating sialyl-3-T antigen (Neu5Acα2-3Galβ1-3GalNAcα-O-Ser/Thr, Figure 1). The sialyl-3-T antigen can be further sialylated by ST6GalNAC-I, II, III resulting in the disialylated-T antigen (Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAcα-O-Ser/Thr; disialyl-T; Figure 1). On the other hand, the T antigen can also experience O-6 sialylation by either ST6GalNAc-I and/or ST6GalNAc-II, preferentially, ST6GalNAc-II, originating sialyl-6-T (Galβ1-3(Neu5Acα2-6)GalNAcα-O-Ser/Thr) [5, 38-40]. The above described low molecular weight O-GalNAc glycans and their sialylated counterparts are often present in tumor tissues as a result of a premature stop in glycosylation. This structures are commonly referred to as either simple mucin type O-glycans or T-related antigens [3].

Core 1 O-GalNAc is usually the starting point of much longer and more complex structures. Namely, the addition of a GlcNAc from UDP-GlcNAc to the O-6 position of the GalNAc residue by the core 2 β1-6-*N*-acetylglucosaminyltransferase (C2GnT) results in the formation of core 2 (GlcNAcβ1-6(Galβ1-3)GalNAcα-O-Ser/Thr; Figure 1 and Table 1) [41]. The C2GnTs present themselves in three isoforms, C2GnT-1, -2 and -3. The C2GnT-1/3, also known as leukocyte type (L type), is expressed in many tissues and synthesizes only the core 2. The C2GnT-2, also known as mucin type (M type) is found only in mucin-secreting cell types. This isoform is responsible for the synthesis of core 2 as

well as core 4 (GlcNAcβ1-6(GlcNAcβ1-3)GalNAcα-O-Ser/Thr; Figure 1 and Table 1) [41, 42]. Conversely, in the presence of core 3 β1-3-*N*-acetylglucosaminyltransferase (C3GnT) the Tn antigen can act as a precursor of core 3 (GlcNAcβ1-3GalNAcα-O-Ser/Thr; Figure 1 and Table 1) that can further originate core 4. The core 4 (GlcNAcβ1-6(GlcNAcβ1-3)GalNAcα-O-Ser/Thr; Figure 1 and Table 1) synthesis is accomplish by the action of M type β1-6-*N*-acetylglucosaminetransferase (C2GnT-2). This enzyme transfer GlcNAc from UDP-GlcNAc to the O-6 position of GalNAc [43]. The core 3 and branched core 4 O-GalNAc glycans have been found only in secreted mucins of certain mucin-secreting tissues, such as intestinal mucus [23] and salivary glands [5, 44].

The extension of the O-GalNAc cores is achieved by the action of β1-4/3 galactosyltranferases (β4/3GalTs) and β1-3/4-*N*-acetylglucosaminyltransferases (β3/4GnTs), Figure 1, which produces poly-lactosamine residues (Galβ1-4GlcNAcβ-), termed type II chains, or poly-neolactosamine residues (Galβ1-3GlcNAcβ-), termed type I chains. The elongated O-glycans can be further terminated by either Lewis and/or ABO blood determinants [45]. The *O*-glycan capping may also occur by the concerted addition of Neu5Ac through sialyltransferases (ST6Gal and ST3Gal), or by other molecules such as sulphate [5, 46].

Table 1. Structures of O-GalNAc glycans cores.

O-glycan cores	Structure
Core 1 or T antigen	Galβ1-3GalNAcα-O-Ser/Thr
Core 2	GlcNAcβ1-6(Galβ1-3)GalNAcα-O-Ser/Thr
Core 3	GlcNAcβ1-3GalNAcα-O-Ser/Thr
Core 4	GlcNAcβ1-6(GlcNAcβ1-3)GalNAcα-O-Ser/Thr
Core 5	GalNAcα1-3GalNAcα-O-Ser/Thr
Core 6	GlcNAcβ1-6GalNAcα-O-Ser/Thr
Core 7	GalNAcα1-6GalNAcα-O-Ser/Thr
Core 8	Galα1-3GalNAcα-O-Ser/Thr

Cores 1-4 are the major cores expressed in human cells. Still, other core structures, such as core structures 5 to 8, have been observed, but are not so common in nature. Mucins with core 5 O-GalNAc glycans (GalNAcα1-3GalNAcα-O-Ser/Thr; Table 1) have been reported in human meconium and intestinal adenocarcinoma tissue, whereas core 6 O-GalNAc (GlcNAcβ1-6GalNAcα-O-Ser/Thr; Table 1) is found in human intestinal mucin and ovarian cyst mucin. Core 8 O-GalNAc glycans (Galα1-3GalNAcα-O-Ser/Thr; Table 1) has been reported in human respiratory mucin, while bovine sub maxillary mucin may express core 7 O-GalNAc glycans (GalNAcα1-6GalNAcα-O-Ser/Thr; Table 1) [23].

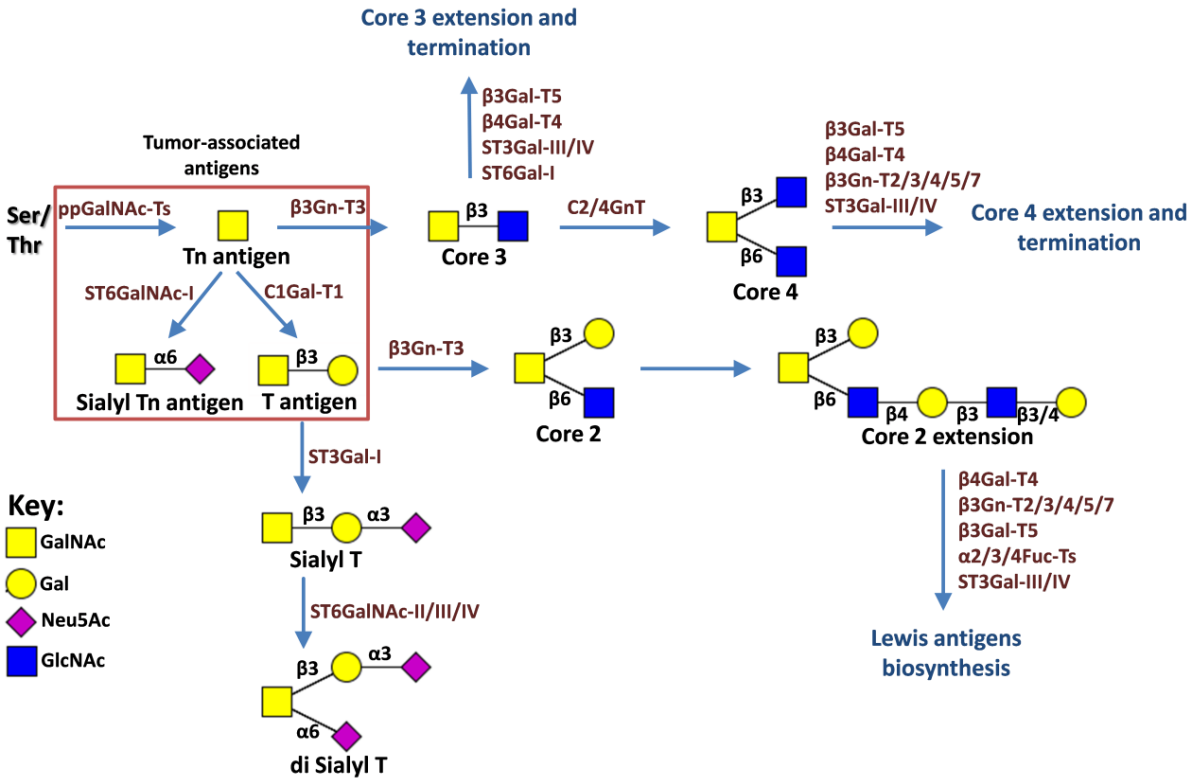


Figure 1. The biosynthetic pathway of O-GalNAc glycosylation and the action of competitive glycosyltransferases: polypeptide *N*-acetylgalactosaminyltransferases (ppGalNAc-Ts); sialyltransferases ST6GalNAc and ST3/ST6Gal; core 1 β1-3 galactosyltransferase (C1GalT); core 3 β1-3 *N*-acetylglucosaminyltransferase (C3GnT); Core 2 β1-6-*N* acetylglucosaminyltransferase (C2GnT); fucosyltransferases (αFuc-Ts), β1-4/3 galactosyltranferases (β4/3Gal-Ts) and β1-3/4 *N*-acetylglucosaminyltransferases (β3/4Gn-Ts).

3. Alterations at O-glycosylation level in tumors

The pattern of protein glycosylation is cell and tissue specific [26] and closely reflects the physiological status of the cell. Thereby, changes in glycan expression are being observed for several pathological conditions and in particular in during malignancy (reviewed by Dabelsteen *et al*. (1996) [47]) .

It is well established that changes in cellular O-glycosylation occur during malignant transformation, however, the functional significance of these alterations is less well understood [48]. Changes in carbohydrate core structure, increased synthesis of terminal structures, including Lewis blood group related antigens have been reported [47]. Indeed, glycan changes in malignant cells may assume a variety of forms, yet, the majority of the studies consensually report a high expression of incomplete O-glycans such as sialyl-Tn and other sialylated TF-related antigens [3]. Among the mechanisms leading to their expression are the up or down regulations of glycosyltransferases, re-localization of the enzymes, disorganization of secretory pathways, and metabolic alterations [5, 26].

The sialyl-Tn is a pancarcinoma antigen which is aberrantly expressed in several epithelial tumors, including colorectal [49], ovarian [50], breast [51], pancreatic [52] and gastric carcinomas [7], while absent in the healthy mucosa [53]. The expression of sialyl-Tn has been mostly connoted with the high expression of sialyltransferases, in particularly ST6GalNAc-I. In accordance with this observations, studies in carcinogenesis reported that this enzyme is weakly expressed in normal gastric mucosa [6]. On the other hand, Marcos *et al (2004)* showed that transfected gastric cells with ST6GalNAc-I expressed high levels of sialyl-Tn and this sialyltransferase is the major responsible for the its synthesis *in vitro* [34]. More recently ST6GalNAc-I was observed as the major enzyme controlling the expression of cancer-associated sialyl-Tn antigen in the gastrointestinal tissues[6].

The high expression of enzymes is not the only mechanism responsible for the (over)expression of sialyl-Tn. Several reports have been described a translocation of ST6GalNAc-I from later to early sub-compartments of the Golgi apparatus, as well as a blockage in the expression of core 1 and core 3. This phenomenon makes the Tn epitope available for sialyltransferases and thus leading the formation of sialyl-Tn. Another mechanism was explored in the genetic field that found mutations in Cosmc gene, that

encodes Cosmc, a molecular chaperone required for the activity of C1GalT. This mutation has been related with an enhanced synthesis of sialyl-Tn and Tn antigens [46, 53, 54].

Given the overexpression or *de novo* expression of sialyl-Tn and its responsible mechanisms, this epitope has been associated with aggressiveness and metastatic potential of the epithelial carcinomas [55]. It is also a independent indicator of poor prognosis in gastric [56], ovarian [57] and colorectal cancer [58]. Regarding the biological value of sialyl-Tn in tumors and its expression at cell surface, therapeutic approaches has been developing cancer drugs against this epitope. Studies reported the detection of naturally occurring antibodies to sialyl-Tn in the serum of gastrointestinal and ovarian cancer patients raising the possibility of immune response towards this epitope [59]. Therefore, several strategies have been purposed for targeting sialyl-Tn on tumor cell surfaces in order to increase the production of antibodies and T cell response [48].

Beyond the sialyl-Tn as a major glycan of interest, the sialyl-T and disialyl-T antigens are commonly observed in healthy tissue, however they have been found overexpressed in several tumors [60]. These include testicular neoplasms [61], bladder tumors [9] and pancreatic adenocarcinomas [62]. The sialyl-T has been associated with worse prognosis and its overexpression is mediated by an increase in the expression of α2-3 sialyltransferases, namely ST3Gal-I [63]. Another study concerning pancreatic carcinoma has demonstrated that sialyl-T and disialyl-T can result from a down regulation in core 2 β1-6-*N*-acetylglucosaminyltransferase, which is required for elongation of the mucin-type glycans [62].

4. Analytical strategies to analyze sialylated TF-related O-glycans by Mass Spectrometry

Due to their biological relevance much research as focused in sialylated TF-related antigens. Moreover, the identification of these modifications in proteins is considered the first step towards understanding structure–function relationships [18]. Classical approaches involve mainly immunoblotting with specific antibodies and/or lectins with or without the combinations of enzymatic treatments to obtain structures of interest. Even though useful in a routine basis, structural characterization is nowadays warranted for validation of novel

targets. In this context, matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) has been proven a powerful, highly sensitive, tool for the structural analysis of complex glycans [18, 64-67].

In MALDI, the energy of the laser is first absorbed by the matrix molecules. The stimulated matrix molecules transfer their excess energy to the sample molecules, which become ionized. For this reason, MALDI is a soft ionization method that produces predominantly single charges ions from the released glycans [22]. Moreover, it has the advantage of sample preparation and a relatively high tolerance to salts [68] and other contaminants, such as buffers [69]. This has enable the screen of complex mixtures for oligosaccharides from biological extracts, thereby revealing the type of glycans present [22]. However, the identification of acidic sugars, such as sialic acids, and deoxysugars, such as fucose, constitutes a major analytical challenge for MALDI instruments. Both during ionization and under typical vacuum source conditions, MALDI induces strong dissociation of labile glycosidic bonds in the analyzer, a process called metastable fragmentation [12, 70]. Derivatization strategies such as permethylation have allowed to stabilize this substituents overcoming the described limitation [71, 72]. However some drawbacks related with the formation of byproducts and the harshness of the procedure has strongly limited the sensitivity of this approach [20, 73]. As such, publications related with the analysis of permethylated sialylated O-glycans by MALDI have shown a limit of detection no greater than 80 *pico*mol [74].

4.1. Methods for the isolation of O-GalNAc glycoproteins

Given the drawbacks associated with the analysis of sialylated TF-related glycans by MALDI, several enrichment methods have been introduced for O-GalNAc glycoproteins. The most popular enrichment techniques include the affinity chromatography using antibodies or lectins to selectively bind complex carbohydrate structures and discriminate between subtly different glycan forms [75, 76]. In addition to the specific lectin-affinity, approaches based on general physical and chemical properties glycoproteins are most valuable [77]. Among them, beads and nanoparticles are been functionalized with molecules with high affinity for the O-glycoproteins of interest [78-80]. Then the purified glycoproteins are separated in a one dimension or two dimension gel electrophoresis [77]. After separation, proteins are transferred by electroblotting onto a

membrane for specific detection. The specific binding of diverse lectins to glycans can be used for detection and isolation of glycoproteins from human body fluids such as serum, and saliva [81]. It can be particularly informative to compare the "profiles" of proteins in a complex sample with those revealed with the specific lectin, indicating which proteins are glycosylated and/or aberrantly glycosylated [82, 83].

After enrichment/isolation of the glycoprotein of interest, O-glycans need to be released from the aglicone. Even though enzymes exist to remove *N*-glycans this is not the case of *O*-linked sugars. As there are no generic enzyme able to cleave O-glycans this is often achieved using chemical methods, in particular alkaline β-elimination [74]. This method was first suggested by Carlson *et al*. (1968) and is, nowadays, considered the most reliable for this purpose [84]. Still, this is approach is known to induce profound degradation of the protein moiety and peeling of the O-glycans. To reduce the degradation of the glycosidic domain the de-O-glycosylation is generally performed in the presence of sodium borohydride [85]. This reducing agent reduces the anomeric carbon (Figure 2) in the reducing end of the released O-glycans thus stopping further degradation [17]. Then the O-glycans can be isolated from other contaminants by affinity chromatography with lectins or antibodies [86], ion exchange [22], filtration and size-exclusion chromatography [87, 88].

4.2. Analysis of sialylated glycans by permethylation

After being released, the oligosaccharides are subjected to per-*O*-methylation for further structural analysis. Permethylation of sialic acids according Ciucanu and Kerek (1984) [89] method has been found most effective for analysis by MALDI MS [71, 72]. Permethylation of sialylated glycans has particular advantages to those underivatitized glycans due to (1) better ionization responses [12, 17, 18], (2) smaller mass increase and a greater volatility [90, 91] and (3) consequently profiled effectively using vacuum MALDI MS and tandem MS methods [12, 92, 93]. Although these advantages must be balanced against the disadvantages of the derivatization step itself. The permethylation chemistry has lower limit on the quantity of starting material, thereby it may be not appropriate for analysis available in small quantities, such as from gel spots [12, 22].

The permethylation procedure starts with the deprotonation of the hydroxyl groups and other proton-donating groups in the presence of a base, such as sodium hydroxide

(NaOH) in anhydrous dimethyl sulfoxide (DMSO). The DMSO is a polar aprotic solvent, with high affinity for both native and permethylated glycans and it does not establish hydrogen bonds. Powdered NaOH is then introduced in the reaction media, which not only ensures the deprotonation of hydrogen donating groups in the glycans but also contributes to highly anhydrous conditions necessary for the reaction. The permethylation is then achieved by the introduction of a strong alkylating reagent such as methyl iodine (Figure 2). Then, the reaction is stopped by the addiction of water and the permethylated glycans are extracted in chloroform or dichloromethane [12, 89, 94].

Using optimal proportion of powdered NaOH, DMSO and methyl iodide the method gives high yields in one step without side products and with no special care to avoid exposure to air and humidity [95]. However, there are oxidative degradation and peeling reactions associated with the high pH resulting from dissolving NaOH powder prior to liquid-liquid extractions. These side reactions are adversely prominent with small samples (low *pico*mole to *fento*mole) [21]. However, simple procedures have been found to improve and avoid the formation of oxidation products. The addiction of the base first and later methyl iodide avoids the salts formed during the interaction of DMSO and methyl iodide[96, 97]; the water that is used to stop the reaction can eliminate the oxidation products as well as neutralization of the base with a strong acid [20].

In attempts to introduce a semi-quantitative approach some studies report the introduction of internal standards such as β-ciclodextrine[18]. For a quantitative purpose studies reported labeling permethylation involving stable isotope label with $^{13}CH_3I$ [98, 99]. Another development in permethylation method involves packing of sodium hydroxide powder in micro-spin columns or fused-silica capillaries, allowing effective derivatization [21]. In this approach, prior to mass spectrometry, analytes are mixed with methyl iodide in a dimethyl sulfoxide solution containing traces of water and infusing through the micro-reactors containing powdered NaOH. These solid phase permethylation as shown to be able to accurately and rapidly analyze *pico*mole amounts of linear and branched, sialylated and neutral glycans samples. These authors further suggest that this type of approach reduce oxidative degradation and peeling reactions and avoid the need of excessive clean-up.

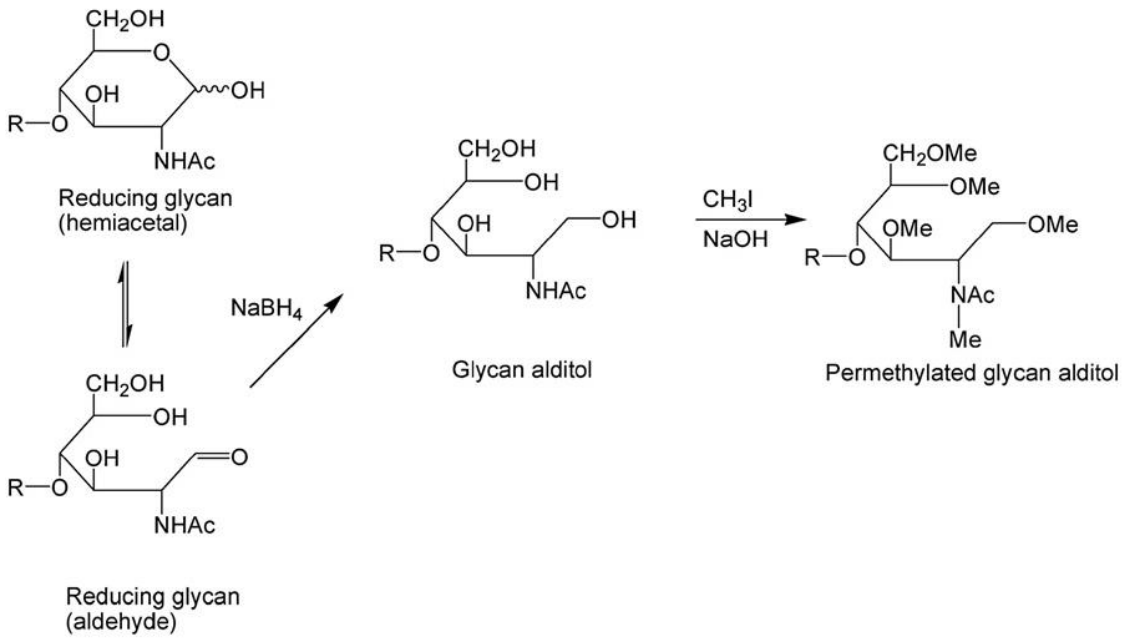


Figure 2. Reduction of oligosaccharides with sodium borohydride and further permethylation. Adpated from [17].

4.3. Collision Induced Dissociation of oligosacharides

Accurate structural validation of O-glycans by mass spectrometry can only be fully achievable by tandem experiments. Collision induced dissociation (CID), is the most commonly used method for fragmenting oligosaccharide ions and it is used in the MALDI TOF/TOF equipment. Ions collide with small neutral molecules to convert the ions kinetic energy to vibrational energy [100]. This process involves first isolating the ion of interest, the precursor ion, from a mixture of ions generated during the ionization event. The ion is excited and collided with an inert gas such as helium, nitrogen, or argon to produce fragments. Kinetic energy from the collision is converted into internal vibrational energy that leads to bond breaking reactions [100, 101]. The process can be repeated multiple times depending on the mass analyzer. There are several parameters that influence the fragmentation behavior of oligosaccharides during the CID event. They include the collision energy, the amount of internal energy deposited in precursor ions upon collisions, the number of collisions, and the time scale between collision activation and detection.

CID methods can be classified in major categories based on the translational energy of the precursor ions, low and high energy [102].

High energy CID results in the excitation of the electronic states in the precursor ion. Several types of fragmentation reactions can occur because high-energy collisions produce a broad internal energy distribution. MALDI TOF/TOF has been shown to produce high-energy collision conditions for oligosaccharides [101]. High energy collision conditions produce larger fractions of cross-ring cleavages such as A and X type ions [103, 104]. However in native oligosaccharides, high energy CID often yields many peaks that are artifacts due to metastable dissociation. Interpreting the spectra of unknown compounds is therefore complicated by the presence of fake peaks [105].

For oligosaccharides, CID can provide sequence, branching [106], and even linkage [107] and stereochemistry [108, 109].Oligosaccharides that contain the same monosaccharides linked with a different branching structure often show distinct product ion patterns because the steric environments differ between such isomers and result in different bond energies and ion abundances in product-ion mass spectra [107].

The fragmentation of oligosaccharide ions is assigned according to the nomenclature introduced by Domon and Costello (1988) [110] (Figure 3), where product ions containing the reducing end of the oligosaccharide are labeled with letters from the end of the alphabet, X (cross-ring cleavage), Y and Z (glycosidic bond cleavage). Those fragments containing the non-reducing end are termed A (cross-ring cleavage), and B and C (glycosidic bond cleavage). Subscript numerals indicate cleavage along the glycosidic bond, whereas superscript numerals denote the position of the cross-ring cleavage (Figure 3).

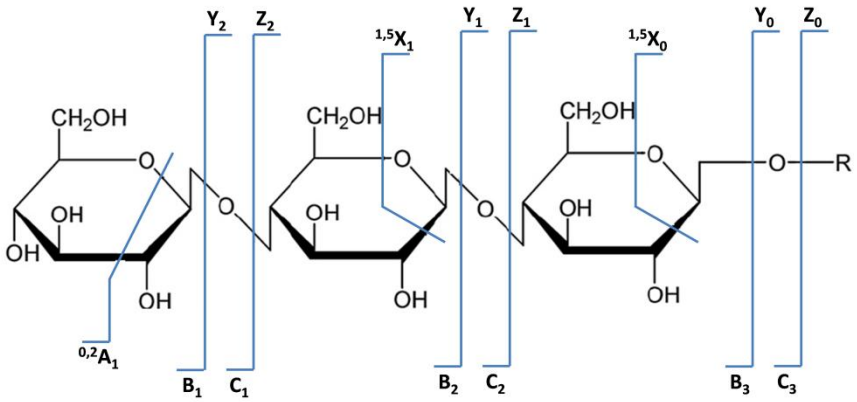


Figure 3. The nomenclature for oligosaccharide fragmentation according Domon and Costello (1988).

4.4. CID of permethylated oligosaccharides

The conversion of glycans to hydrophobic derivatives has also advantages on the tandem MS of permethylated oligosaccharides relative to those in native forms. When the permethylated glycans are subjected to CID, the glycan bond scission occurring during tandem MS origins "scars". They are product ions with unique masses by virtue of the lack of methyl or methoxy groups, that indicate the topology and linkages for glycans [13]. On the other hand, tandem mass spectra of native are less structural informative, because when a glycosidic bond cleavage occurs, the product ion masses do not determine which bond was cleaved (glycosidic bond cleavages do not leave a mass "scar") [18].

The permethylated derivatives are often ionized as $[M+Na]^+$ ions unless reversed phase LC/MS is used, under which conditions protonated ions ($[M+H]^+$) may also be formed [92, 93, 107]. The fragmentation energies of the protonated ions from permethylated glycans are lower than the sodiated ones [111]. Therefore product ion patterns of protonated and sodiated forms contain different features. Thus, in the protonated ions from permethylated glycans produces mostly ions containing the non reducing end [12], while the sodiated ion fragments form abundant reducing and non-reducing terminal product ions with approximately equal abundances [112]. Despite this main differences observed, both permethylated forms have in common several characteristics, (1) the preferential cleavage in the reducing side of HexNAc residues [91, 113] (2) formation of a quaternary ammonium cation from the free amine groups that is differentiated by mass from the *N*-acetyl-hexosamine (HexNAc) residues [114]; and (3) the permethylated derivatives produces useful multiple bond fragmentation mass spectra [115]. This is an important advantage, because the *m/z* values of such ions are, in several situations, not distinguishable from those of primary fragments ions generated from underivatized oligosaccharides [17].

Relative to the CID spectra from the protonated permethylated glycans exhibits an additional ion type, the E ion. The E ion comes from eliminative losses of substituents to the 3-position of HexNAc that is very useful for the identification of the type of inter glycosidic linkage [116]. Nevertheless, in the fragmentation of protonated permethylated ions may undergoes by internal residue rearrangement, which difficulties the spectrum analysis .

Taking in account the characteristics of the product ion spectra from permethylated glycans, the tandem MS of sialylated permethylated glycans is well established approach to determine overall branching structure and some structural isomers [19]. The product ion spectra of sialylated permethylated glycans exhibits in abundant Neu5Ac oxonium, which complicates the location position of substitution, because ions with glycosidic bonds to those residues are in low abundance [12]. Despite this, Lemoine *et al.* (1991) described that it is possible to differentiate sialic acid linkage isomers, α2-3 and α2-6, through CID fragmentation, due to specific A type ions occurring to the galactose residue to which Neu5Ac is attached. The permethylated isomer with Neu5Acα2-6 linked produces a pattern in which the most abundant product ions contain the non reducing terminus and those that contain the reducing terminus are in low abundance. On other hand the permethylated isomer Neu5Acα2-3Hexose produces abundant product ion containing the reducing terminus [117]. The CID product ion mass spectra of sodiated permethylated oligosaccharides is used instead protonated ones because provide additional sensitivity to the linkage structure [12].

5. References

1. Apweiler, R., H. Hermjakob, and N. Sharon, *On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database.* Biochim Biophys Acta, 1999. **1473**(1): p. 4-8.
2. Ohtsubo, K. and J.D. Marth, *Glycosylation in cellular mechanisms of health and disease.* Cell, 2006. **126**(5): p. 855-67.
3. Reis, C.A., H. Osorio, L. Silva, C. Gomes, and L. David, *Alterations in glycosylation as biomarkers for cancer detection.* J Clin Pathol, 2010. **63**(4): p. 322-9.
4. Kim, Y.S., J. Gum, Jr., and I. Brockhausen, *Mucin glycoproteins in neoplasia.* Glycoconj J, 1996. **13**(5): p. 693-707.
5. Brockhausen, I., *Pathways of O-glycan biosynthesis in cancer cells.* Biochim Biophys Acta, 1999. **1473**(1): p. 67-95.
6. Marcos, N.T., E.P. Bennett, J. Gomes, A. Magalhaes, C. Gomes, L. David, I. Dar, C. Jeanneau, S. DeFrees, D. Krustrup, L.K. Vogel, E.H. Kure, J. Burchell, J. Taylor-Papadimitriou, H. Clausen, U. Mandel, and C.A. Reis, *ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues.* Front Biosci (Elite Ed), 2011. **3**: p. 1443-55.

7. David, L., J.M. Nesland, H. Clausen, F. Carneiro, and M. Sobrinho-Simoes, *Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases.* APMIS Suppl, 1992. **27**: p. 162-72.
8. MacLean, G.D., M. Reddish, R.R. Koganty, T. Wong, S. Gandhi, M. Smolenski, J. Samuel, J.M. Nabholtz, and B.M. Longenecker, *Immunization of breast cancer patients using a synthetic sialyl-Tn glycoconjugate plus Detox adjuvant.* Cancer Immunol Immunother, 1993. **36**(4): p. 215-22.
9. Videira, P.A., M. Correia, N. Malagolini, H.J. Crespo, D. Ligeiro, F.M. Calais, H. Trindade, and F. Dall'Olio, *ST3Gal.I sialyltransferase relevance in bladder cancer tissues and cell lines.* BMC Cancer, 2009. **9**: p. 357.
10. Heimburg-Molinaro, J., M. Lum, G. Vijay, M. Jain, A. Almogren, and K. Rittenhouse-Olson, *Cancer vaccines and carbohydrate epitopes.* Vaccine, 2011. **29**(48): p. 8802-26.
11. Bernardes, G.J., B. Castagner, and P.H. Seeberger, *Combined approaches to the synthesis and study of glycoproteins.* ACS Chem Biol, 2009. **4**(9): p. 703-13.
12. Zaia, J., *Mass spectrometry of oligosaccharides.* Mass Spectrom Rev, 2004. **23**(3): p. 161-227.
13. Reinhold, V.N., B.B. Reinhold, and C.E. Costello, *Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID.* Anal Chem, 1995. **67**(11): p. 1772-84.
14. Moyer, S.C. and R.J. Cotter, *Atmospheric pressure MALDI.* Anal Chem, 2002. **74**(17): p. 468A-476A.
15. O'Connor, P.B. and C.E. Costello, *A high pressure matrix-assisted laser desorption/ionization Fourier transform mass spectrometry ion source for thermal stabilization of labile biomolecules.* Rapid Commun Mass Spectrom, 2001. **15**(19): p. 1862-8.
16. Leymarie, N. and J. Zaia, *Effective Use of Mass Spectrometry for Glycan and Glycopeptide Structural Analysis.* Analytical Chemistry, 2012. **84**(7): p. 3040-3048.
17. Zaia, J., *Mass spectrometry and the emerging field of glycomics.* Chem Biol, 2008. **15**(9): p. 881-92.
18. Zaia, J., *Mass spectrometry and glycomics.* OMICS, 2010. **14**(4): p. 401-18.
19. Mechref, Y., P. Kang, and M.V. Novotny, *Differentiating structural isomers of sialylated glycans by matrix-assisted laser desorption/ionization time-of-flight/time-of-flight tandem mass spectrometry.* Rapid Commun Mass Spectrom, 2006. **20**(8): p. 1381-9.
20. Ciucanu, I. and C.E. Costello, *Elimination of oxidative degradation during the per-O-methylation of carbohydrates.* J Am Chem Soc, 2003. **125**(52): p. 16213-9.
21. Kang, P., Y. Mechref, I. Klouckova, and M.V. Novotny, *Solid-phase permethylation of glycans for mass spectrometric analysis.* Rapid Commun Mass Spectrom, 2005. **19**(23): p. 3421-8.
22. Morelle, W., V. Faid, F. Chirat, and J.C. Michalski, *Analysis of N- and O-linked glycans from glycoproteins using MALDI-TOF mass spectrometry.* Methods Mol Biol, 2009. **534**: p. 5-21.
23. Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1908/

24. Clausen, H. and E.P. Bennett, *A family of UDP-GalNAc: Polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation.* Glycobiology, 1996. **6**(6): p. 635-646.
25. Ten Hagen, K.G., T.A. Fritz, and L.A. Tabak, *All in the family: the UDP-GalNAc : polypeptide N-acetylgalactosaminyltransferases.* Glycobiology, 2003. **13**(1): p. 1R-16R.
26. Gill, D.J., H. Clausen, and F. Bard, *Location, location, location: new insights into O-GalNAc protein glycosylation.* Trends Cell Biol, 2011. **21**(3): p. 149-58.
27. Mandel, U., H. Hassan, M.H. Therkildsen, J. Rygaard, M.H. Jakobsen, B.R. Juhl, E. Dabelsteen, and H. Clausen, *Expression of polypeptide GalNAc-transferases in stratified epithelia and squamous cell carcinomas: immunohistological evaluation using monoclonal antibodies to three members of the GalNAc-transferase family.* Glycobiology, 1999. **9**(1): p. 43-52.
28. Hassan, H., C.A. Reis, E.P. Bennett, E. Mirgorodskaya, P. Roepstorff, M.A. Hollingsworth, J. Burchell, J. Taylor-Papadimitriou, and H. Clausen, *The lectin domain of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities.* J Biol Chem, 2000. **275**(49): p. 38197-205.
29. Wandall, H.H., H. Hassan, E. Mirgorodskaya, A.K. Kristensen, P. Roepstorff, E.P. Bennett, P.A. Nielsen, M.A. Hollingsworth, J. Burchell, J. Taylor-Papadimitriou, and H. Clausen, *Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3.* J Biol Chem, 1997. **272**(38): p. 23503-14.
30. Perrine, C.L., A. Ganguli, P. Wu, C.R. Bertozzi, T.A. Fritz, J. Raman, L.A. Tabak, and T.A. Gerken, *Glycopeptide-preferring polypeptide GalNAc transferase 10 (ppGalNAc T10), involved in mucin-type O-glycosylation, has a unique GalNAc-O-Ser/Thr-binding site in its catalytic domain not found in ppGalNAc T1 or T2.* J Biol Chem, 2009. **284**(30): p. 20387-97.
31. Tian, E. and K.G. Ten Hagen, *Recent insights into the biological roles of mucin-type O-glycosylation.* Glycoconj J, 2009. **26**(3): p. 325-34.
32. Kato, K., H. Takeuchi, A. Kanoh, U. Mandel, H. Hassan, H. Clausen, and T. Irimura, *N-acetylgalactosamine incorporation into a peptide containing consecutive threonine residues by UDP-N-acetyl-D-galactosaminide:polypeptide N-acetylgalactosaminyltransferases.* Glycobiology, 2001. **11**(10): p. 821-9.
33. Ikehara, Y., N. Kojima, N. Kurosawa, T. Kudo, M. Kono, S. Nishihara, S. Issiki, K. Morozumi, S. Itzkowitz, T. Tsuda, S.I. Nishimura, S. Tsuji, and H. Narimatsu, *Cloning and expression of a human gene encoding an N-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens.* Glycobiology, 1999. **9**(11): p. 1213-24.
34. Marcos, N.T., S. Pinho, C. Grandela, A. Cruz, B. Samyn-Petit, A. Harduin-Lepers, R. Almeida, F. Silva, V. Morais, J. Costa, J. Kihlberg, H. Clausen, and C.A. Reis, *Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen.* Cancer Res, 2004. **64**(19): p. 7050-7.
35. Ju, T. and R.D. Cummings, *A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16613-8.

36. Ju, T., R.P. Aryal, C.J. Stowell, and R.D. Cummings, *Regulation of protein O-glycosylation by the endoplasmic reticulum-localized molecular chaperone Cosmc.* J Cell Biol, 2008. **182**(3): p. 531-42.
37. Takashima, S., Y. Tachida, T. Nakagawa, T. Hamamoto, and S. Tsuji, *Quantitative analysis of expression of mouse sialyltransferase genes by competitive PCR.* Biochem Biophys Res Commun, 1999. **260**(1): p. 23-7.
38. Brockhausen, I., J. Yang, M. Lehotay, S. Ogata, and S. Itzkowitz, *Pathways of mucin O-glycosylation in normal and malignant rat colonic epithelial cells reveal a mechanism for cancer-associated Sialyl-Tn antigen expression.* Biol Chem, 2001. **382**(2): p. 219-32.
39. Lee, Y.C., M. Kaufmann, S. Kitazume-Kawaguchi, M. Kono, S. Takashima, N. Kurosawa, H. Liu, H. Pircher, and S. Tsuji, *Molecular cloning and functional expression of two members of mouse NeuAcalpha2,3Galbeta1,3GalNAc GalNAcalpha2,6-sialyltransferase family, ST6GalNAc III and IV.* J Biol Chem, 1999. **274**(17): p. 11958-67.
40. Kono, M., T. Tsuda, S. Ogata, S. Takashima, H. Liu, T. Hamamoto, S.H. Itzkowitz, S. Nishimura, and S. Tsuji, *Redefined substrate specificity of ST6GalNAc II: a second candidate sialyl-Tn synthase.* Biochem Biophys Res Commun, 2000. **272**(1): p. 94-7.
41. Bierhuizen, M.F. and M. Fukuda, *Expression cloning of a cDNA encoding UDP-GlcNAc:Gal beta 1-3-GalNAc-R (GlcNAc to GalNAc) beta 1-6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen.* Proc Natl Acad Sci U S A, 1992. **89**(19): p. 9326-330.
42. Bierhuizen, M.F., K. Maemura, and M. Fukuda, *Expression of a differentiation antigen and poly-N-acetyllactosaminyl O-glycans directed by a cloned core 2 beta-1,6-N-acetylglucosaminyltransferase.* J Biol Chem, 1994. **269**(6): p. 4473-9.
43. Kuhns, W., V. Rutz, H. Paulsen, K.L. Matta, M.A. Baker, M. Barner, M. Granovsky, and I. Brockhausen, *Processing O-glycan core 1, Gal beta 1-3GalNAc alpha-R. Specificities of core 2, UDP-GlcNAc: Gal beta 1-3 GalNAc-R(GlcNAc to GalNAc) beta 6-N-acetylglucosaminyltransferase and CMP-sialic acid: Gal beta 1-3GalNAc-R alpha 3-sialyltransferase.* Glycoconj J, 1993. **10**(5): p. 381-94.
44. Brockhausen, I., K.L. Matta, J. Orr, and H. Schachter, *Mucin synthesis. UDP-GlcNAc:GalNAc-R beta 3-N-acetylglucosaminyltransferase and UDP-GlcNAc:GlcNAc beta 1-3GalNAc-R (GlcNAc to GalNAc) beta 6-N-acetylglucosaminyltransferase from pig and rat colon mucosa.* Biochemistry, 1985. **24**(8): p. 1866-74.
45. Brooks, S.A., T.M. Carter, L. Royle, D.J. Harvey, S.A. Fry, C. Kinch, R.A. Dwek, and P.M. Rudd, *Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies.* Anticancer Agents Med Chem, 2008. **8**(1): p. 2-21.
46. Bard, F., D.J. Gill, and H. Clausen, *Location, location, location: new insights into O-GaINAc protein glycosylation.* Trends in Cell Biology, 2011. **21**(3): p. 149-158.
47. Dabelsteen, E., *Cell surface carbohydrates as prognostic markers in human carcinomas.* J Pathol, 1996. **179**(4): p. 358-69.
48. Hakomori, S., *Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines.* Adv Exp Med Biol, 2001. **491**: p. 369-402.

49. Itzkowitz, S.H., E.J. Bloom, W.A. Kokal, G. Modin, S. Hakomori, and Y.S. Kim, *Sialosyl-Tn - a Novel Mucin Antigen Associated with Prognosis in Colorectal-Cancer Patients.* Cancer, 1990. **66**(9): p. 1960-1966.
50. Kobayashi, H., T. Terao, and Y. Kawashima, *Serum sialyl Tn as an independent predictor of poor prognosis in patients with epithelial ovarian cancer.* J Clin Oncol, 1992. **10**(1): p. 95-101.
51. Yonezawa, S., T. Tachikawa, S. Shin, and E. Sato, *Sialosyl-Tn antigen. Its distribution in normal human tissues and expression in adenocarcinomas.* Am J Clin Pathol, 1992. **98**(2): p. 167-74.
52. Kim, G.E., H.I. Bae, H.U. Park, S.F. Kuan, S.C. Crawley, J.J. Ho, and Y.S. Kim, *Aberrant expression of MUC5AC and MUC6 gastric mucins and sialyl Tn antigen in intraepithelial neoplasms of the pancreas.* Gastroenterology, 2002. **123**(4): p. 1052-60.
53. Ju, T., G.S. Lanneau, T. Gautam, Y. Wang, B. Xia, S.R. Stowell, M.T. Willard, W. Wang, J.Y. Xia, R.E. Zuna, Z. Laszik, D.M. Benbrook, M.H. Hanigan, and R.D. Cummings, *Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc.* Cancer Res, 2008. **68**(6): p. 1636-46.
54. Wang, Y., T. Ju, X. Ding, B. Xia, W. Wang, L. Xia, M. He, and R.D. Cummings, *Cosmc is an essential chaperone for correct protein O-glycosylation.* Proc Natl Acad Sci U S A, 2010. **107**(20): p. 9228-33.
55. Springer, G.F., *Tn epitope (N-acetyl-D-galactosamine alpha-O-serine/threonine) density in primary breast carcinoma: a functional predictor of aggressiveness.* Mol Immunol, 1989. **26**(1): p. 1-5.
56. Werther, J.L., S. Rivera-MacMurray, H. Bruckner, M. Tatematsu, and S.H. Itzkowitz, *Mucin-associated sialosyl-Tn antigen expression in gastric cancer correlates with an adverse outcome.* Br J Cancer, 1994. **69**(3): p. 613-6.
57. Holmberg, L.A., D.V. Oparin, T. Gooley, K. Lilleby, W. Bensinger, M.A. Reddish, G.D. MacLean, B.M. Longenecker, and B.M. Sandmaier, *Clinical outcome of breast and ovarian cancer patients treated with high-dose chemotherapy, autologous stem cell rescue and THERATOPE STn-KLH cancer vaccine.* Bone Marrow Transplant, 2000. **25**(12): p. 1233-41.
58. Itzkowitz, S.H., E.J. Bloom, T.S. Lau, and Y.S. Kim, *Mucin associated Tn and sialosyl-Tn antigen expression in colorectal polyps.* Gut, 1992. **33**(4): p. 518-23.
59. Holmberg, L.A. and B.M. Sandmaier, *Theratope vaccine (STn-KLH).* Expert Opin Biol Ther, 2001. **1**(5): p. 881-91.
60. Hakomori, S., *Tumor-associated carbohydrate antigens.* Annu Rev Immunol, 1984. **2**: p. 103-26.
61. Rajpert-De Meyts, E., S.N. Poll, I. Goukasian, C. Jeanneau, A.S. Herlihy, E.P. Bennett, N.E. Skakkebaek, H. Clausen, A. Giwercman, and U. Mandel, *Changes in the profile of simple mucin-type O-glycans and polypeptide GalNAc-transferases in human testis and testicular neoplasms are associated with germ cell maturation and tumour differentiation.* Virchows Arch, 2007. **451**(4): p. 805-14.
62. Beum, P.V., J. Singh, M. Burdick, M.A. Hollingsworth, and P.W. Cheng, *Expression of core 2 beta-1,6-N-acetylglucosaminyltransferase in a human pancreatic cancer cell line results in altered expression of MUC1 tumor-associated epitopes.* J Biol Chem, 1999. **274**(35): p. 24641-8.
63. Gillespie, W., J.C. Paulson, S. Kelm, M. Pang, and L.G. Baum, *Regulation of alpha 2,3-sialyltransferase expression correlates with conversion of peanut agglutinin*

(PNA)+ to PNA- phenotype in developing thymocytes. J Biol Chem, 1993. **268**(6): p. 3801-4.
64. Kyselova, Z., Y. Mechref, P. Kang, J.A. Goetz, L.E. Dobrolecki, G.W. Sledge, L. Schnaper, R.J. Hickey, L.H. Malkas, and M.V. Novotny, *Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles.* Clin Chem, 2008. **54**(7): p. 1166-75.
65. de Leoz, M.L., L.J. Young, H.J. An, S.R. Kronewitter, J. Kim, S. Miyamoto, A.D. Borowsky, H.K. Chew, and C.B. Lebrilla, *High-mannose glycans are elevated during breast cancer progression.* Mol Cell Proteomics, 2011. **10**(1): p. M110 002717.
66. Tang, Z., R.S. Varghese, S. Bekesova, C.A. Loffredo, M.A. Hamid, Z. Kyselova, Y. Mechref, M.V. Novotny, R. Goldman, and H.W. Ressom, *Identification of N-glycan serum markers associated with hepatocellular carcinoma from mass spectrometry data.* J Proteome Res, 2010. **9**(1): p. 104-12.
67. Williams, T.I., D.A. Saggese, K.L. Toups, J.L. Frahm, H.J. An, B. Li, C.B. Lebrilla, and D.C. Muddiman, *Investigations with O-linked protein glycosylations by matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry.* J Mass Spectrom, 2008. **43**(9): p. 1215-23.
68. Beavis, R.C. and B.T. Chait, *High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry.* Anal Chem, 1990. **62**(17): p. 1836-40.
69. Finke, B., M. Mank, H. Daniel, and B. Stahl, *Offline coupling of low-pressure anion-exchange chromatography with MALDI-MS to determine the elution order of human milk oligosaccharides.* Anal Biochem, 2000. **284**(2): p. 256-65.
70. Wada, Y., P. Azadi, C.E. Costello, A. Dell, R.A. Dwek, H. Geyer, R. Geyer, K. Kakehi, N.G. Karlsson, K. Kato, N. Kawasaki, K.H. Khoo, S. Kim, A. Kondo, E. Lattova, Y. Mechref, E. Miyoshi, K. Nakamura, H. Narimatsu, M.V. Novotny, N.H. Packer, H. Perreault, J. Peter-Katalinic, G. Pohlentz, V.N. Reinhold, P.M. Rudd, A. Suzuki, and N. Taniguchi, *Comparison of the methods for profiling glycoprotein glycans - HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study.* Glycobiology, 2007. **17**(4): p. 411-422.
71. Viseux, N., C.E. Costello, and B. Domon, *Post-source decay mass spectrometry: optimized calibration procedure and structural characterization of permethylated oligosaccharides.* J Mass Spectrom, 1999. **34**(4): p. 364-76.
72. Juhasz, P. and C.E. Costello, *Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass-Spectrometry of Underivatized and Permethylated Gangliosides.* Journal of the American Society for Mass Spectrometry, 1992. **3**(8): p. 785-796.
73. York, W.S., L.L. Kiefer, P. Albersheim, and A.G. Darvill, *Oxidation of Oligoglycosyl Alditols during Methylation Catalyzed by Sodium-Hydroxide and Iodomethane in Methyl Sulfoxide.* Carbohydrate Research, 1990. **208**: p. 175-182.
74. Goetz, J.A., M.V. Novotny, and Y. Mechref, *Enzymatic/Chemical Release of O-Glycans Allowing MS Analysis at High Sensitivity.* Analytical Chemistry, 2009. **81**(23): p. 9546-9552.
75. Wu, A.M., E. Lisowska, M. Duk, and Z. Yang, *Lectins as tools in glycoconjugate research.* Glycoconj J, 2009. **26**(8): p. 899-913.
76. Dai, Z., J. Zhou, S.J. Qiu, Y.K. Liu, and J. Fan, *Lectin-based glycoproteomics to explore and analyze hepatocellular carcinoma-related glycoprotein markers.* Electrophoresis, 2009. **30**(17): p. 2957-66.

77. Wuhrer, M., M.I. Catalina, A.M. Deelder, and C.H. Hokke, *Glycoproteomics based on tandem mass spectrometry of glycopeptides.* J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **849**(1-2): p. 115-28.
78. Abbott, K.L., K. Aoki, J.M. Lim, M. Porterfield, R. Johnson, R.M. O'Regan, L. Wells, M. Tiemeyer, and M. Pierce, *Targeted glycoproteomic identification of biomarkers for human breast carcinoma.* J Proteome Res, 2008. **7**(4): p. 1470-80.
79. Abbott, K.L., J.M. Lim, L. Wells, B.B. Benigno, J.F. McDonald, and M. Pierce, *Identification of candidate biomarkers with cancer-specific glycosylation in the tissue and serum of endometrioid ovarian cancer patients by glycoproteomic analysis.* Proteomics, 2010. **10**(3): p. 470-81.
80. Tang, J., Y. Liu, P. Yin, G.P. Yao, G.Q. Yan, C.H. Deng, and X.M. Zhang, *Concanavalin A-immobilized magnetic nanoparticles for selective enrichment of glycoproteins and application to glycoproteomics in hepatocelluar carcinoma cell line.* Proteomics, 2010. **10**(10): p. 2000-2014.
81. Gilboa-Garber, N., B. Lerrer, E. Lesman-Movshovich, and O. Dgani, *Differential staining of Western blots of human secreted glycoproteins from serum, milk, saliva, and seminal fluid using lectins displaying diverse sugar specificities.* Electrophoresis, 2005. **26**(23): p. 4396-401.
82. Hashii, N., N. Kawasaki, S. Itoh, M. Hyuga, T. Kawanishi, and T. Hayakawa, *Glycomic/glycoproteomic analysis by liquid chromatography/mass spectrometry: analysis of glycan structural alteration in cells.* Proteomics, 2005. **5**(18): p. 4665-72.
83. Kim, Y.S., S.Y. Hwang, S. Oh, H. Sohn, H.Y. Kang, J.H. Lee, E.W. Cho, J.Y. Kim, J.S. Yoo, N.S. Kim, C.H. Kim, E. Miyoshi, N. Taniguchi, and J.H. Ko, *Identification of target proteins of N-acetylglucosaminyl-transferase V and fucosyltransferase 8 in human gastric tissues by glycomic approach.* Proteomics, 2004. **4**(11): p. 3353-8.
84. Carlson, D.M., *Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins.* J Biol Chem, 1968. **243**(3): p. 616-26.
85. Mechref, Y., *Analysis of glycans derived from glycoconjugates by capillary electrophoresis-mass spectrometry.* Electrophoresis, 2011. **32**(24): p. 3467-81.
86. Wuhrer, M., A.M. Deelder, and C.H. Hokke, *Protein glycosylation analysis by liquid chromatography-mass spectrometry.* Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2005. **825**(2): p. 124-133.
87. Brooks, S.A., *Strategies for analysis of the glycosylation of proteins: current status and future perspectives.* Mol Biotechnol, 2009. **43**(1): p. 76-88.
88. Geyer, H. and R. Geyer, *Strategies for analysis of glycoprotein glycosylation.* Biochim Biophys Acta, 2006. **1764**(12): p. 1853-69.
89. Ciucanu, I. and F. Kerek, *A Simple and Rapid Method for the Permethylation of Carbohydrates.* Carbohydrate Research, 1984. **131**(2): p. 209-217.
90. McNeil, M., A.G. Darvill, P. Aman, L.E. Franzen, and P. Albersheim, *Structural analysis of complex carbohydrates using high-performance liquid chromatography, gas chromatography, and mass spectrometry.* Methods Enzymol, 1982. **83**: p. 3-45.
91. Dell, A., H.R. Morris, H. Egge, H. Vonnicolai, and G. Strecker, *Fast-Atom-Bombardment Mass-Spectrometry for Carbohydrate-Structure Determination.* Carbohydrate Research, 1983. **115**(Apr): p. 41-52.

92. Viseux, N., E. de Hoffmann, and B. Domon, *Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry.* Analytical Chemistry, 1998. **70**(23): p. 4951-4959.
93. Viseux, N., E. deHoffmann, and B. Domon, *Structural analysis of permethylated oligosaccharides by electrospray tandem mass spectrometry.* Analytical Chemistry, 1997. **69**(16): p. 3193-3198.
94. Ciucanu, I., *Per-O-methylation reaction for structural analysis of carbohydrates by mass spectrometry.* Anal Chim Acta, 2006. **576**(2): p. 147-55.
95. Larson, G., H. Karlsson, G.C. Hansson, and W. Pimlott, *Application of a simple methylation procedure for the analyses of glycosphingolipids.* Carbohydr Res, 1987. **161**(2): p. 281-90.
96. Johnson, C.R. and W.G. Phillips, *Reactions of Alkoxysulfonium Salts with Alkoxides.* Journal of Organic Chemistry, 1967. **32**(6): p. 1926-&.
97. Paul, R.C., P. Singh, and S.L. Chadha, *Chemistry of Dimethyl Sulphoxide .3. Acid-Base Titrations in Dimethyl Sulphoxide.* Indian Journal of Chemistry, 1970. **8**(11): p. 1010-&.
98. Alvarez-Manilla, G., J. Atwood, 3rd, Y. Guo, N.L. Warren, R. Orlando, and M. Pierce, *Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites.* J Proteome Res, 2006. **5**(3): p. 701-8.
99. Aoki, K., M. Perlman, J.M. Lim, R. Cantu, L. Wells, and M. Tiemeyer, *Dynamic developmental elaboration of N-linked glycan complexity in the Drosophila melanogaster embryo.* J Biol Chem, 2007. **282**(12): p. 9127-42.
100. Medzihradszky, K.F., J.M. Campbell, M.A. Baldwin, A.M. Falick, P. Juhasz, M.L. Vestal, and A.L. Burlingame, *The Characteristics of Peptide Collision-Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer.* Analytical Chemistry, 1999. **72**(3): p. 552-558.
101. Li, B., H.J. An, J.L. Hedrick, and C.B. Lebrilla, *Collision-induced dissociation tandem mass spectrometry for structural elucidation of glycans.* Methods Mol Biol, 2009. **534**: p. 133-45.
102. Li, B., H.J. An, J.L. Hedrick, and C.B. Lebrilla, *Infrared multiphoton dissociation mass spectrometry for structural elucidation of oligosaccharides.* Methods Mol Biol, 2009. **534**: p. 23-35.
103. Mechref, Y., M.V. Novotny, and C. Krishnan, *Structural characterization of oligosaccharides using MALDI-TOF/TOF tandem mass spectrometry.* Anal Chem, 2003. **75**(18): p. 4895-903.
104. Stephens, E., S.L. Maslen, L.G. Green, and D.H. Williams, *Fragmentation characteristics of neutral N-linked glycans using a MALDI-TOF/TOF tandem mass spectrometer.* Anal Chem, 2004. **76**(8): p. 2343-54.
105. An, H.J. and C.B. Lebrilla, *Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial).* Mass Spectrom Rev, 2011. **30**(4): p. 560-78.
106. Domon, B. and C.E. Costello, *Structure elucidation of glycosphingolipids and gangliosides using high-performance tandem mass spectrometry.* Biochemistry, 1988. **27**(5): p. 1534-43.
107. Laine, R.A., K.M. Pamidimukkala, A.D. French, R.W. Hall, S.A. Abbas, R.K. Jain, and K.L. Matta, *Linkage Position in Oligosaccharides by Fast Atom Bombardment Ionization, Collision-Activated Dissociation, Tandem Mass-Spectrometry and*

Molecular Modeling - L-Fucosylp-(Alpha-1-]3)-D-N-Acetyl-D-Glucosamylp-(Beta-1-]3)-D-Galactosylp-(Beta-1-O-Methyl), L-Fucosylp-(Alpha-1-]4)-D-N-Acetyl-D-Glucosaminylp-(Beta-1-]3)-D-Galactosylp-(Beta-1-O-Methyl), L-Fucosylp-(Alpha-1-]6)-D-N-Acetyl-D-Glucosaminylp-(Beta-1-]3)-D-Galactosylp-(Beta-1-O-Methyl)-. Journal of the American Chemical Society, 1988. **110**(21): p. 6931-6939.

108. Dell, A. and H.R. Morris, *Glycoprotein structure determination by mass spectrometry.* Science, 2001. **291**(5512): p. 2351-6.
109. Tseng, K., J.L. Hedrick, and C.B. Lebrilla, *Catalog-library approach for the rapid and sensitive structural elucidation of oligosaccharides.* Anal Chem, 1999. **71**(17): p. 3747-54.
110. Domon, B. and C.E. Costello, *A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-Ms Ms Spectra of Glycoconjugates.* Glycoconjugate Journal, 1988. **5**(4): p. 397-409.
111. Cancilla, M.T., S.G. Penn, J.A. Carroll, and C.B. Lebrilla, *Coordination of alkali metals to oligosaccharides dictates fragmentation behavior in matrix assisted laser desorption ionization Fourier transform mass spectrometry.* Journal of the American Chemical Society, 1996. **118**(28): p. 6736-6745.
112. Egge, H., A. Dell, and H. Von Nicolai, *Fucose containing oligosaccharides from human milk. I. Separation and identification of new constituents.* Arch Biochem Biophys, 1983. **224**(1): p. 235-53.
113. Egge, H. and J. Peterkatalinic, *Fast-Atom-Bombardment Mass-Spectrometry for Structural Elucidation of Glycoconjugates.* Mass Spectrometry Reviews, 1987. **6**(3): p. 331-393.
114. Baldwin, M.A., N. Stahl, L.G. Reinders, B.W. Gibson, S.B. Prusiner, and A.L. Burlingame, *Permethylation and tandem mass spectrometry of oligosaccharides having free hexosamine: analysis of the glycoinositol phospholipid anchor glycan from the scrapie prion protein.* Anal Biochem, 1990. **191**(1): p. 174-82.
115. Richter, W.J., D.R. Muller, and B. Domon, *Tandem mass spectrometry in structural characterization of oligosaccharide residues in glycoconjugates.* Methods Enzymol, 1990. **193**: p. 607-23.
116. Weiskopf, A.S., P. Vouros, and D.J. Harvey, *Electrospray ionization-ion trap mass spectrometry for structural analysis of complex N-linked glycoprotein oligosaccharides.* Anal Chem, 1998. **70**(20): p. 4441-7.
117. Lemoine, J., G. Strecker, Y. Leroy, B. Fournet, and G. Ricart, *Collisional-activation tandem mass spectrometry of sodium adduct ions of methylated oligosaccharides: sequence analysis and discrimination between alpha-NeuAc-(2----3) and alpha-NeuAc-(2----6) linkages.* Carbohydr Res, 1991. **221**: p. 209-17.
118. Ceroni, A., K. Maass, H. Geyer, R. Geyer, A. Dell, and S.M. Haslam, *GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans.* J Proteome Res, 2008. **7**(4): p. 1650-9.

Chapter II. Aims and scopes

O-glycosylation and aberrantly O-glycosylated proteins play a major role in tumor biology, biomarker research and therapeutics. The present work is within the scope of collaboration between Portuguese Institute of Oncology (IPO) of Porto and (Institute of Molecular Pathology and Immunology of the University of Porto) IPATIMUP. The purpose of this work was to develop an analytical approach for the analysis of trace amounts of sialylated TF-related antigens by MALDI TOF/TOF from low *pico*mole of glycoproteins isolated by SDS-PAGE. The first part of the work is devoted to the establishment of the analytical bases to achieve this goal, while the second part has been directed towards the application the technology. Particular attention has been devoted to the detection of sialyl-Tn in glycoproteins from bladder and gastric cancer precursor lesions isolated from urine and serum, respectively.

6. Analytical Approaches

The sialylated TF-related antigens, namely sialyl-Tn, sialyl-T and disialyl-T, were analyzed by permethylation of O-glycans followed by analysis by nano-HPLC-MALDI TOF/TOF. Briefly glycoproteins were removed from the gel and de-O-glycosylated by reductive β-elimination at 45ºC for 16 hours. After the released glycan were permethylated based on an adaptation of the protocol described by Ciucanu and Kerek (1984) [89]. The permethylated O-glycans were filtred with a cutoff at 10KDa and a removal of borates was accomplished with 5% acetic acid in methanol. Then the resulting glycans were separated by nano-HPLC on a C18 reverse phase column, collected on a MALDI plate and analyzed by MALDI-TOF/TOF using a 2,5-dihydroxybenzoic acid (DHB) matrix [22]. A bioinformatics survey of the ions of interest as used to enhance the selection of the relevant analytical signals for MS/MS. Ion assignments were made based on the MS spectra and confirmed with the corresponding fragmentations by MS/MS. Spectra interpretation was done using the glycoworkbench platform [118]. This procedure was established using fetuin, a glycoprotein known to express sialylated TF-related antigens. The mentioned analytical platform was further applied to validate the expression of sialylated-TF antigens in glycoproteins from bladder and gastric cancer precursor lesions isolated from urine and serum, respectively. Whenever necessary, affinity chromatography methods were applied to recover the proteins of interest from their complex milieus.

Chapter III. Detection of sialylated cancer-associated Thomsen-Friedenreich - related antigens by in gel deglycosylation and nanoLC-MALDI-TOF/TOF

ABSTRACT

Malignant transformations are accompanied by an (over)expression of sialylated Thomsen-Friedenreich -related antigens, namely sialyl-Tn, sialyl-T and disialyl-T. These epitopes result from a premature stop in protein O-glycosylation and have direct impact on tumor biology. The proteins carrying these alterations can ultimately be shed into the blood stream, urine and other body fluids and thus be explored as biomarkers in non-invasive tests. Currently their detection relies on immuno-based methods that, even though useful in a routine basis, often fail to fully highlight the glycosylation pattern of a given protein. Herein, we have systematized a target-driven analytical approach to determine these glycans in minute amounts of glycoproteins (*pico*mole) isolated from SDS-PAGE gels. Briefly, the proteins are to be de-O-glycosylated in gel by reductive beta elimination, permethylated and analyzed by nanoLC-MALDI-TOF/TOF with enhanced selection of the relevant analytical signals using bioinformatics tools. This has been established using fetuin as a model protein and used to validate the expression of sialyl-Tn in a glycoprotein isolated from the urine of rats with chemically-induced bladder tumors. To our knowledge, the presence of bladder tumor-associated sialyl-Tn antigens in urine is being reported for the first time and more studies are ongoing to evaluate the clinical value of this observation. The described approach is regarded to be of primary interest to laboratories equipped with nanoLC-MALDI-TOF/TOF equipments and involved in glycobiology research.

Keywords: tumor glycosylation, mucin-type O-glycans, cancer biomarkers, Thomsen-Friedenreich antigen, bladder cancer, MALDI

INTRODUCTION

Thomsen-Friedenreich(TF)-related antigens are a class of low molecular weight O-glycans that include the T antigen (Galβ1–3GalNAcα-O-Ser/Thr) and its precursor Tn (GalNAcα-O-Ser/Thr). In most human healthy tissues the T antigen is further extended and is often terminated by Lewis and ABO blood group antigens. Alternatively, Tn and T antigens can be sialylated by sialyltransferases forming sialyl-Tn (Neu5Acα2-6GalNAcα-O-Ser/Thr), sialyl-3-T (Neu5Acα2-3Galβ1-3GalNacα-O-Ser/Thr), its positional isomer sialyl-6-T (Galβ1–3(NeuAcα2–6)GalNAcα-O-Ser/Thr), and disialyl-T (NeuAcα2–3Galβ1–3(NeuAcα2–6)GalNAcα-O-Ser/Thr), thereby prematurely stopping the elongation of the O-glycan chain (reviewed in Reis *et al*. 2010) [1].

The *de novo* expression of sialylated TF-related antigens is characteristic of several human malignancies [2]. These abnormal changes in protein glycosylation have been mostly attributed to a disorganization of secretory organelles and/or altered glycosyltransferases expression. Frequently, an increased expression of a given glycoprotein amplifies the alterations in secretory pathways resulting in a more pronounced accumulation of otherwise absent sialylated antigens. Moreover, it may contribute to enhancement in the basal levels of antigens already present in healthy tissues. In particular, sialyl-Tn has been found in breast [3], esophagus [4], colon[5], pancreas [6], stomach [7], endometrium [8], bladder [9], and ovary [10] carcinomas, whereas low or no expression is observed in the respective normal tissues [11]. The presence of this epitope contributes to alter cell recognition by the immune system [12], affects the adhesive properties of cancer cells, promotes invasion and metastasis[7, 13, 14]. As a result, it is being explored as a serum biomarker of poor prognosis in gastric [15, 16], colorectal [17] and ovarian cancer carcinomas and in tumor vaccines [18]. Sialyl-T and disialyl-T have also been found overexpressed in several cancers, namely testicular neoplasms [19], bladder tumors [20] and pancreatic adenocarcinomas [21]. Given their role in tumor behavior as well as biomarker and therapeutic values, much effort has been put into the identification of tumor-associated proteins carrying sialylated TF-related antigens.

The sialylated TF-related antigens have been classically determined based on immunoblotting [22] and more recently by *in situ* proximity ligation [23]. Even though useful in a routine basis, these methods require further characterization of the glycosylation pattern of a given glycoprotein. Also, validation assays using structural determination are

required, particularly when dealing with unknown and/or highly complex matrices. Moreover, at the moment there are no commercially available antibodies for mono- and disialylated TF-antigens. Thus the most common strategy to access these antigens is based on their reactivity to antibodies and lectins such as peanut agglutinin (PNA) targeting the T antigen, after neuraminidase treatment. As a result these sialylated glycans have been classified under the general designation of cryptic T antigens, which fails to provide information about their particular structural nature.

In this context, matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) has been proven a powerful, highly sensitive, tool for the structural analysis of complex glycans [24-28]. However, MALDI ionization induces a high degree of vibrational excitation in the formed ions leading to a significant dissociation of labile substituents [29, 30]. The analysis by mass spectrometry of sialic acid-containing glycans is particularly challenging, as they experience prompt fragmentation and decomposition [28, 31]. Permethylation overcomes this limitation by stabilizing these labile substituents [28], allowing their detection in positive ion mode [32]. By equalizing their chemical properties it also permits a semi-quantitative comparison between the permethylated glycans [33]. Permethylated glycans are typically observed as sodiated ions, unless reversed phase LC/MS is used, under such conditions protonated ions are also formed [34, 35]. Tandem MS of permethylated glycans is also known to produce more structural detail than their native counterparts [36]. High energy collision induction dissociation (CID) of sodium ions produces mostly cross ring cleavages (A and X types; nomenclature by Domon and Costello 1988 [37]). In particular, the masses of specific A-type ions in Gal residues substituted with Neu5Ac have allowed the distinction between α2-3 and α2-6 linkages in modern MALDI TOF/TOF instruments [38]. Conversely, protonated species exhibit mostly ions corresponding to the rupture of glycosidic bonds (B, C, Y types). In both cases useful product ion patterns are formed, that ultimately have allowed the assignment of isomeric structures [39, 40]. Furthermore, the MS/MS of protonated, permethylated glycans does not exhibit misleading transfer products observable for native glycans [41]. Despite these advantages, permethylation-based analysis is conventionally considered not satisfactory at the range of low *pico*mole to *femto*mole quantities, which are usually the amounts at which glycoproteins are present in biological samples [42]. This

limitation has hampered the use of permethylation as a viable option for the analysis of glycans released from SDS-PAGE gel spots.

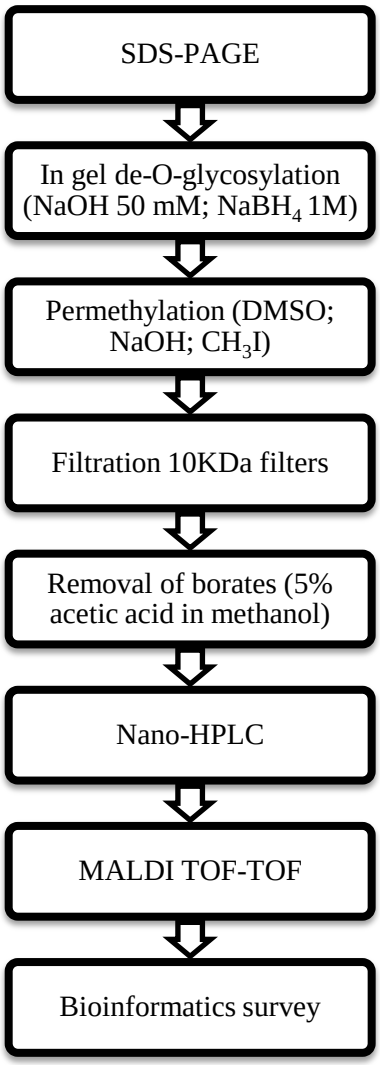
Herein, we have demonstrate that a permethylation-based nanoHPLC-MALDI-TOF-TOF approach can be used to validate and complement information from western blotting regarding the expression of sialylated Thomsen-Friedenreich-related antigens. This has been proven successful using as starting material minute amounts (low *pico*mol range) of proteins isolated by SDS-PAGE.

EXPERIMENTAL

Unless otherwise stated, all reagents were purchased from Sigma-Aldrich.

Proof of concept

Fetuin from fetal calf serum was used as model glycoproteins in this study. Aliquots of 0.1 (approximately 1.6 *pico*mol), 0.25, 0.5, 1, 1.5 µg (approximately 24 *pico*mol) of fetuin were electrophoresed in 4-12% Mini-PROTEAN TGX Precast Gels (BIORAD). Fetuin glycans were then recovered from the gel and analyzed by MALDI-TOF/TOF as summarized in Scheme 1. Briefly, the two main bands observed after staining with Coomassie Brilliant Blue G-250 (Pierce-Thermo Scientific) were subsequently removed and subjected to in gel de-O-glycosylation. The release O-glycans were then permethylated and analyzed by nanoLC-MALDI-TOF/TOF with enhanced selection of the relevant analytical signals using bioinformatics tools (Scheme 1). Equivalent amounts of fetuin previously digested with neuraminidase from *Clostridium perfringens* were used as controls. After validation, the analytical approach presented in Scheme 1 was applied to the analysis of bladder tumor associated urinary glycoproteins. Preliminary studies were also performed to determine the limit of detection of MALDI-TOF/TOF for permethylated O-glycans released by chemical de-glycosylation from native fetuin and purified by filtration. These studies included fetuin amounts between 50 ng (approximately 0.8 *pico*mol) and 1.5 µg (approximately 24 *pico*mol. All experiments were done in triplicates.



Scheme 1. Overview on the analytical approach developed for the analysis of trace amounts of sialylated TF-related glycans.

Animal model

Twenty female Wistar rats were obtained at the age of 5 weeks from Harlan (Amsterdam, Netherlands). The rats were used in this study after a week of acclimatization. During the experimental protocol, animals were housed in collective plastic cages with *ad libitum* access to food and water. The room temperature and the relative humidity were controlled at 22±2ºC and 60±10%, respectively. Fluorescent lighting was provided in a 12-h light/dark cycle. The animals were randomly divided into two groups: control group (n=10 rats) and urothelial carcinogenesis group induced by exposure to *N*-butyl-*N*-(4-hydroxybutyl)nitrosamine (BBN; n=10 rats) described by

Oliveira *et al*. (2007) [43]. All procedures were performed in accordance with the European Communities Council Directive 2010/63/EU .

Induction and characterization of BBN-resulting urothelial carcinogenesis

In order to induce urothelial carcinogenesis, one group of animals (BBN; Tokyo Kasei Kogyo) was treated with BBN. The BBN was administered in the drinking water, in light impermeable bottles, at a concentration of 0.05%. Control animals served as the control without any chemical supplementation in drinking water. The animals were observed daily for health check. The urothelial carcinogenesis group was exposed to BBN for 20 weeks and was maintained with normal tap water until the end of the experiment. After 28 weeks, the animals from control group and urothelial carcinogenesis group were both were sacrificed with 0.4% sodium pentobarbital (1 ml/Kg, intraperitoneal) and gastrocnemius muscle was dissected out. Complete necropsies were carefully conducted. All organs were examined macroscopically for any changes. The urinary bladders were inflated in *situ* by injection of 10% phosphate-buffered formalin (300 µl), ligated around the neck to maintain proper distension and then were immersed in the same solution for 12 h. After fixation, the formalin was removed; the urinary bladder was weighed and cut into two parts and was routinely processed for haematoxylin and eosin staining.

Immunohistochemistry

Formalin-fixed paraffin embedded (FFPE) urothelium sections from controls and animals with BBN-induced urothelial carcinomas were screened for sialyl-Tn by immunohistochemistry using the avidin/biotin peroxidase method. Briefly, 2 µm sections were deparaffinised and hydrated with xylene and graded ethanol series, and exposed to 3% hydrogen peroxide in methanol for 20 minutes, to reduce endogenous peroxidase activity. No antigen retrieval steps were necessary to determine this particular epitope. The expression of sialyl-Tn was then evaluated using anti-sialyl-Tn mouse monoclonal antibody TKH2. The sections were first blocked with BSA (5% in PBS) for 30 minutes to avoid non-specific and then incubated with TKH2 at room temperature for 30 min. Afterwards the sections were washed with PBS-T and incubated at room temperature for 30 minutes with the Vectastain Elite ABC peroxidase kit (Vector Lab). After washing in

PBS-T, the sections were incubated with 3,3-diaminobenzidine tetrahydrochloride (DAB, Dako) for 5 minutes at room temperature to visualize antibody binding sites. Finally, they were counterstained with hematoxylin for 1 minute. Positive and negative control sections of intestinal metaplasias were tested in parallel. The negative control sections were performed by adding BSA (5% in PBS) devoid of any antibody. The expression of this epitope was evaluated by microscopic screening of sections exhibiting a maximum of negative and positive stained cells.

Urine collection and isolation of urine glycoproteins

The urines from control group and urothelial carcinogenesis group were collected for 24 hours using metabolic cages. The urines were then centrifuged at 8,000 *g* at room temperature for 15 min to remove cells and debris and conserved at -80ºC. No special treatment was performed to remove occasional microhematuria, in order to preserve the intrinsic components of the sample. After thawing, each sample was dialyzed with Spectra/Por MWCO 12-14 (Spectrum Laboratories) against dionized water, lyophilized, redissolved in 5% SDS, dialyzed again against 1.5% SDS at 60ºC and finally dionized water as described by Halim *et al.* (2012) [44]. The protein content in each sample was determined colorimetry using the RC-DC protein quantification kit (BIORAD). Pools of equal amounts of protein (40 µg) were then constructed for the controls and sialyl-Tn positive tumors using the urine of 3 animals.

Western blotting

Equivalent amounts urinary proteins (20 µg) were separated by gel electrophoresis (12% Mini-PROTEAN TGX Precast Gels (BIORAD)), which were prepared in duplicate for each experiment. The proteins from one of the gels were blotted on a nitrocellulose membrane (Hybond-ECL; Amersham Pharmacia Biotech), and the other gel was stained with Colloidal Coomassie Blue for protein visualization. To verify an equal amount of sample loaded, membranes containing transferred proteins were reversibly stained with Ponceau S. Nonspecific binding to the membrane was blocked with 1% carbo-free blocking solution (*v/v*) (Vector Laboratories). Sialyl-Tn was determined with mouse monoclonal antibody TKH2. Incubation was carried out at 4ºC for 16 hours. Next,

membranes were washed with TBS-T (TBS with 0.5 % Tween 20), incubated with secondary horseradish peroxidase-conjugated anti-mouse or anti-rabbit (GE Healthcare) for 30 min, and washed again with TBS-T. Reactive bands were detected by enhanced chemiluminescence ECL plus (Amersham Pharmacia Biotech) according to the manufacturer's procedure. Images were recorded using X-ray films (Kodak Biomax light Film, Sigma). The films were scanned in Molecular Imager Gel Doc XR+ System (Bio-Rad) and analyzed with QuantityOne software (v 4.6.3 Bio-Rad). Bovine fetuin and albumin were used as positive and negative controls, respectively.

In gel O-deglycosylation and permethylation

The bands of interest were removed from SDS-PAGE gels and de-O-glycosylated by reductive β-elimination upon incubation with 50 mM NaOH and 1M $NaBH_4$ at 45ºC for 16 hours. The reaction was stopped with glacial acetic acid until no fizzing was observed and the samples were subsequently filtered using 10 kDa molecular weight cutoff (MWCO) (Millipore). The filtrated containing low molecular weight peptides, O-glycans and borate salts was recovered and incubated several times with methanol containing 5% (*v/v*) acetic acid under a stream of nitrogen to remove borates as methyl esters.

The O-glycans were then permethylated adopting a modification of the method by Ciucanu and kerek 1988 [45]. Briefly, the native O-glycans were dissolved in 100 µL of DMSO, and powdered NaOH containing trace amounts of water was added to the reaction medium. The mixture was sonicated for 30 min and frozen prior to the addition of 10 µL of CH_3I and then incubated under mild stirring for 1 h. The permethylated oligosaccharide were recovered from the reaction mixture by extraction with dichloromethane and extensively washed with acidified water (pH 2.0) to avoid base-induced hydrolysis of permethylated sialic acids. The samples were then de-salted using Dowex ion-exchange resin (Dowex 50W-X8; biorad).

Nano-HPLC- MALDI-TOF/TOF

The permethylated O-glycans were separated in a nano-HPLC Ultimate 3000 system (Dionex, Amsterdam) equipped with a capillary column (Pepmap100 C18; 3 µm particle size, 0.75 µm internal diameter, 15 cm in length). The samples were dissolved in

5% acetonitrile (ACN) aqueous solution containing 0.1% trifluoroacetic acid (phase A). The separation was performed using a linear gradient of 32-50 % B for 45 minutes, 50-70 % B for 10 minutes and 70-32% A for 5 minutes. The eluted glycans were applied directly on a MALDI plate in 10 seconds fractions using an automatic fraction collector Probot (Dionex, Amsterdam) under a continuous flow rate of 270nL of 2,5-dihydroxybenzoic acid (DHB matrix solution (10 mg/mL in 70% acetonitrile/0.1% TFA and internal standard Glu-Fib at 15ftmol). Mass spectra were obtained on a matrix assisted laser desorption/ionization–time of flight MALDI TOF/TOF mass spectrometer (4800 Proteomics Analyzer, AB SCIEX, Foster City, CA, USA) in the positive ion reflector mode and obtained in the mass range from 600-4500 Da with 1200 laser shots. For the experiment, Glu-Fib was used for internal calibrations. A data dependent acquisition method was created to select the 16 most intense peaks in each sample spot for subsequent tandem mass spectrometry (MS/MS) data acquisition. The MALDI-MS data from each chromatographic run was combined into a three dimensional data array (fraction number, *m/z*, ion current). A survey of plausible analytical signals at *m/z* 669.38, 691.36, 873.48, 895.46, 1234.65 and 1256.64 was performed by determining the most prominent peaks occurring within 0.2 and 0.5 Da of the reference peak. Visual analysis of the distribution of the identified signals allowed the identification of chromatographic envelops. The subsequent MS/MS analysis was governed by the results of this data mining technique and was able to successfully identify the targeted structures. Peak assignments in MS and MS/MS spectra and database searches were performed using the GlycoWorkBench platform [46].

RESULTS AND DISCUSSION

The presence of sialylated TF-related antigens in glycoproteins has been conventionally determined by immunoblotting using antibodies and lectins; still structural characterization is nowadays warranted for validation.

The detection of sialylated glycans by MALDI poses a major analytical challenge due to the significant degree of dissociation of sialic acids under this type of ionization, in particular in the reflector mode [36, 47]. Even though the instruction of permethylation has overcome this limitation by the high pH used in the reaction induces a significant degree of oxidative degradation and peeling reactions. Undesirable by-products are also promoted by the high pH, reductive environment and the temperature necessary for the de-*O*-glycosylation of proteins in the absence of specific O-glycosidases. This synergism of factors is responsible by the introduction of significant background noise and, in some extent, the degradation of the glycans under analysis. As a result, doubts persist about the efficiency of combined approach based upon chemical de-O-glycosylation and permethylation to determined glycosylation patterns from low *pico*mole to *femto*mole quantities of proteins isolated by SDS-PAGE. This work is now focused on the demonstration that in-gel de-O-glycosylation and permethylation coupled with nanoLC-MALDI-TOF/TOF and comprehensive bioinformatics (summarize in Scheme 1) can allow the detection of low *pico*mole of sialylated T-related antigens.

❖ *Determination of sialylated TF-related antigens in fetuin by MALDI*

Fetuin, recognized to yield the two isomeric forms of sialyl-T (s3T and s6T) [48, 49], disialyl-T [50] and, in less extent, sialyl-Tn, was elected as a model glycoprotein for proof of concept. Preliminary studies were performed to determine the limit of detection of MALDI-TOF/TOF for permethylated O-glycans released by chemical de-glycosylation from native fetuin and purified by filtration. This allowed the detection of ions consistent with O-glycans from fetuin amounts as low as 0.1 µg (approximately 1.6 *pico*mol). Based on these observations, amounts of fetuin ranging from 1.6 to 24 *pico*mol were separated in SDS-PAGE and stained with Colloidal Coomassie Blue thereafter. The two main bands observed between 50 and 75 kDa (Figure 1a) were then recovered and subjected to in-gel

de-O-deglycosylated by reductive *β*-elimination. The resulting glycans were isolated by filtration and permethylated in a non-anhydrous environment, as described by Ciucanu and Kerek (1984) [45], to enhance reactions yields and minimize undesirable reactions. The glycans permethylated in this way were further analyzed by MALDI-TOF resulting in MS spectra similar to the one presented in Figure 1b for 4 *pico*mol of fetuin. In accordance with previous reports [48, 49, 51], we have identified several sodium and oxonium ions consistent with O-glycans exhibiting a core 2 backbone (Galβ1-3(Galβ1-4GlcNAcβ1-6)GalNAcα1-Ser/Thr). These assignments have been comprehensively resumed in Table 1 and Figure 1b and include several O-glycans that have not been previously described. Ions belonging to complex type N-glycans were also detected (Table 1, Figure 1b), most likely as consequence of the non-specific nature of the deglycosylation protocol. However, sialylated TF-related glycans could not be accurately distinguished from background noise under the used conditions (Figure 1c).

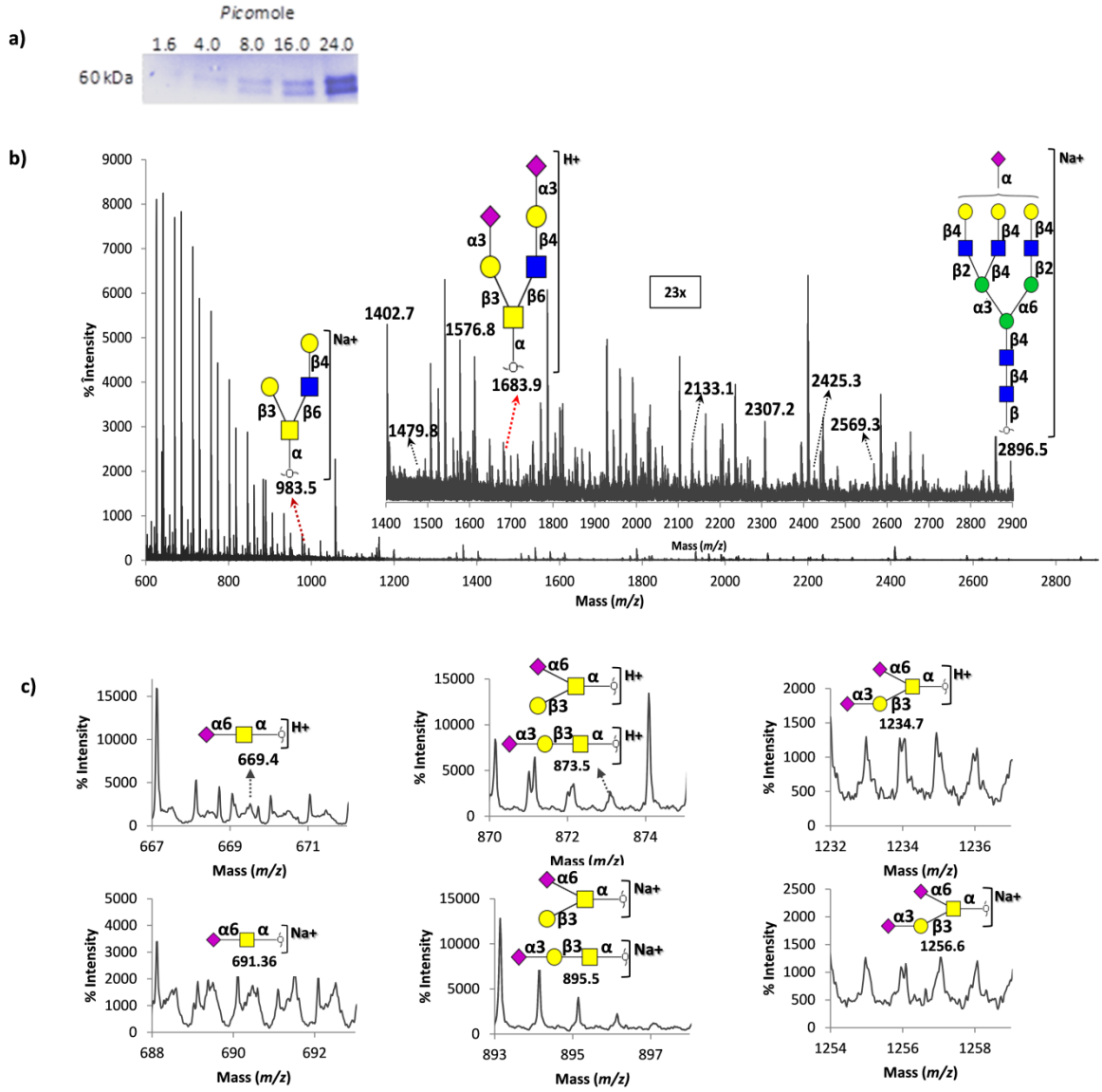
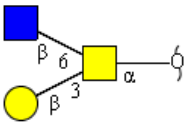
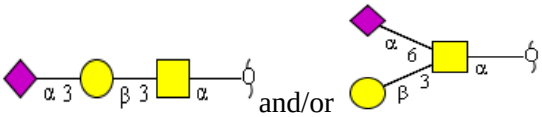
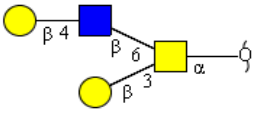
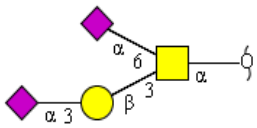
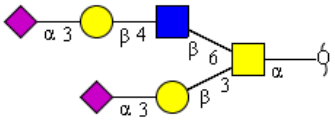
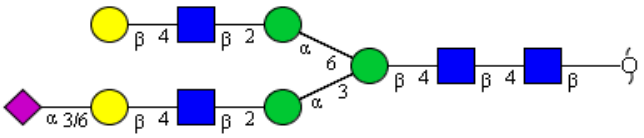
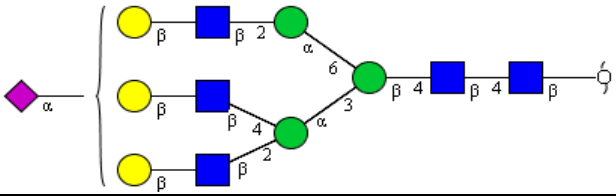


Figure 1. a) Electrophoretic profile of fetuin in 4-12% SDS-PAGE highlighting the amount of glycoprotein used for in gel de-O-glycosylation, b) Positive MALDI-MS spectra of permethylated O-glycans released from 4 *pico*mol of fetuin with an expansion highlighting the ions assigned in Table 1, and c) expansion of the zones of the spectrum were sialyl-Tn, sialyl-T (s-3-T and s-6-T) and disialyl-T ions are expected. (▲ - Fuc; ● - Man; ○ - Gal; □ - GalNAc; ■ - GlcNAc; ◆ - Neu5Ac)

Table 1. Permethylated glycans generated by chemical de-O-glycosylation of 4 *pico*mol of fetuin showing MALDI-MS ions with a signal-to-noise ratio superior to 10. (- Man; - Gal; - GalNAc; - GlcNAc; - Neu5Ac)

ion (*m/z*)	Ion type	Glycosidic family	Assigned structure	Reference
O-glycans				
669.4/691.4*	H^+/Na^+	Core 1		[50]
757.4/779.4	H^+/Na^+	Core 2		[51]
873.5/895.5	H^+/Na^+	Core 2	 and/or 	[48, 49, 51]
983.5	Na^+	Core 2		[49, 51]
1024.5	Na^+	Core 2	$GlcNAc_2Gal_1GalNAcol$	NR
1176.6	H^+	Core 2	$GalNAc_1GlcNAc_1Gal_1GalNAcol$	NR
1228.6	Na^+	Core2/Core 1	$GlcNAc_2Gal_2GalNAcol$	NR
1234.7/1256.6	Na^+	Core 2		[48, 49, 51]
1350.7	H^+	Core 2	$Fuc_2GlcNAc_2Gal_1GalNAcol$	NR
1380.7	H^+	Core 2	$Fuc_1GlcNAc_2Gal_2GalNAcol$	NR
1385.7	Na^+	Core 4	$Neu5Ac_1Gal_1GlcNAc_2GalNAcol$	NR
1402.7	H^+	Core 2	$Fuc_1GlcNAc_2Gal_1GalNAcol$	NR
1479.8	H^+	Core 2	$Neu5Ac_2Gal_1GlcNAc_1GalNAcol$	NR
1535.8	Na^+	Core 2	$Fuc_2GlcNAc_1Gal_3GalNAcol$	NR
1576.8	Na^+	Core 2	$Fuc_2GlcNAc_2Gal_2GalNAcol$	NR
1606.8	Na^+	Core 2	$Fuc_1GlcNAc_2Gal_3GalNAcol$	NR
1683.9	H^+	Core 2		[48, 49, 51]
1933.0	H^+	Core 2	$Fuc_3GlcNAc_2Gal_3GalNAcol$	NR
1987.0	H^+	Core 2	$Neu5Ac_1Gal_2GalNAc_2GalNAcol$	NR
2133.1	H^+	Core 2	$Neu5Ac_2Gal_3GlcNAc_2GalNAcol$	NR
2307.2	H^+	Core 2	$Neu5Ac_2Fuc_1Gal_3GlcNAc_2GalNAcol$	NR
2569.3	H^+	Core 2	$Neu5Ac_1Fuc_2Gal_4GlcNAc_3GalNAcol$	NR

N-glycans				
2425.3	H^+	complex type		[52]
2786.4	H^+	complex type	$Neu5Ac_2Gal_2GlcNAc_3Man_3GlcNAcol$	[52]
2896.5	Na^+	complex type		[52]

* - Traces
NR - non-reported

In an attempt to enhance the signals of interest, the permethylated glycans were then separated by nano-HPLC using a C18 reverse phase column prior to analysis by MALDI. The MALDI-MS data from each chromatographic run was combined into a three dimensional data array (Fraction Number *vs m/z vs* nano-HPLC-MALDI-MS signal) and the distribution of the signals within 0.5 Da of the ions of interest were comprehensively analyzed. The surveyed ions include both protonated and sodiated forms of sialyl-Tn (*m/z* 669.4 and 691.4), sialyl-T (*m/z* 873.5 and 895.7) and disialyl-T (*m/z* 1234.7 and 1256.6). All the above mentioned ions exhibited sparse matrix plots with clusters of high-intensity signals similar to those presented in Figure 2a and b for sialyl-T. These patterns showed a lesser dispersion of *m/z* values over particular regions consistent with chromatographic envelopes (Figure 2a and 2b), strongly suggesting the presence of the targeted compounds.

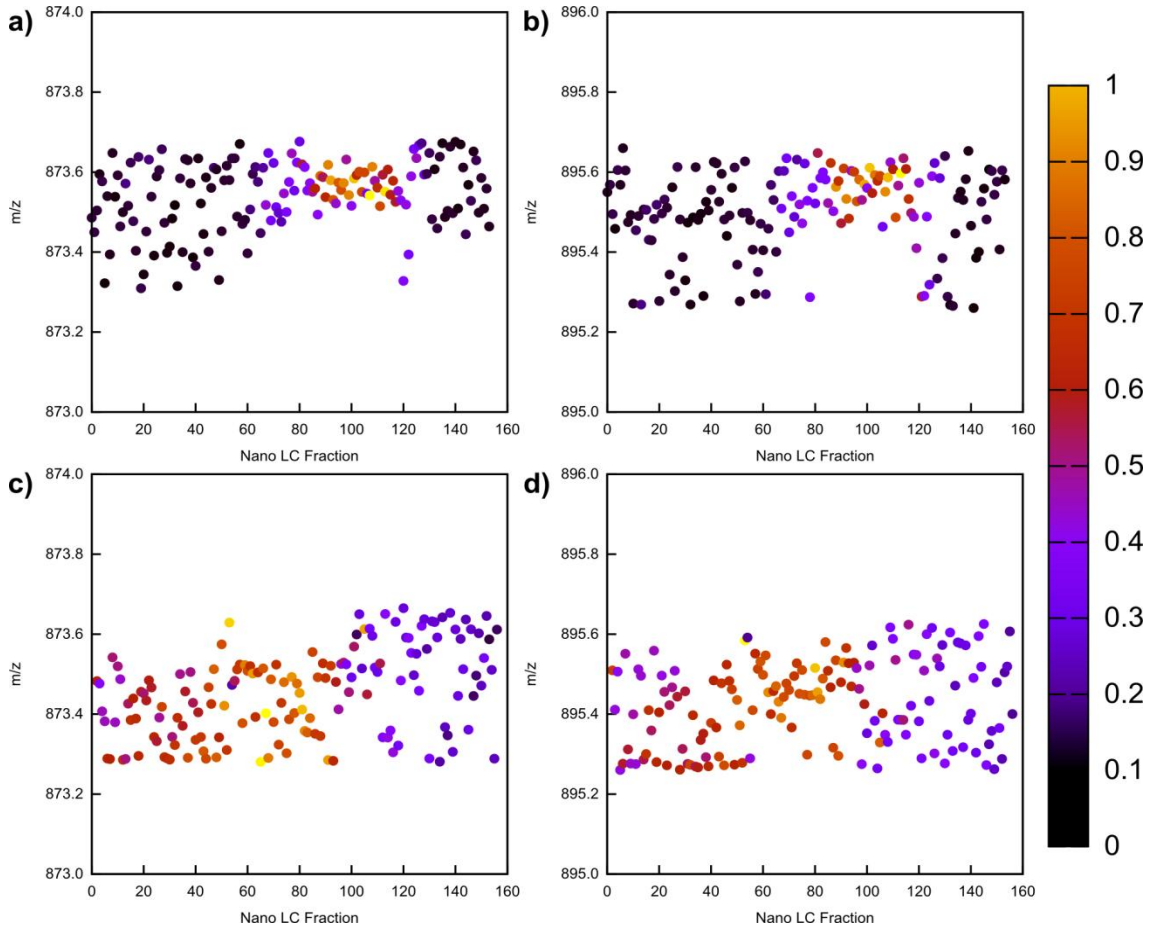


Figure 2. Plots of the sparse matrix (Fraction Number *vs m/z vs* nano-HPLC-MALDI-MS signal) for the peaks found within 873.5±0.5 Da (a and c) and 895.5±0.5Da (b and d), which includes the monoisotopic mass of protonated and sodiated forms of sialyl-T, respectively. Plots a) and b) refer to O-glycans resulting from 4 *pico*mol of native fetuin, whereas c) and d) refer to fetuin subjected to treatment with an α-neuraminidase. A chromatographic envelop is visually detected in figures a) and b) both due to the existence of a cluster of high-intensity signals, but also by the lesser dispersion of *m/z* over the envelope region.

Noteworthy, despite the exposure of the permethylated glycans to a cation-exchange resin prior to analysis and the use of acidified solvents for the chromatographic run, both protonated and sodiated ions were detected. The presence of sodiated and protonated forms most likely results from differences in the coordination of the proton and the sodium cation with the oxygens of the permethylated carbohydrates. According to Cancilla *et al.* (1996) [53] the protonation occurs *via* the oxygen of the glycosidic bond, whereas the sodium ion establishes a more stable binding with multiple oxygen atoms. Based on the generated plots it was possible to accurately identify LC fractions like the

ones presented in Figure 3, showing distinguishable signals for both protonated and sodiated ions. However, the chromatographic envelopes were only present in the assays concerning fetuin contents higher than 1.6 *pico*mol. For amounts of fetuin bellow this range, the poor quality of the spectra did not allow to undoubtedly assign ions associated with permethylated O-glycans.

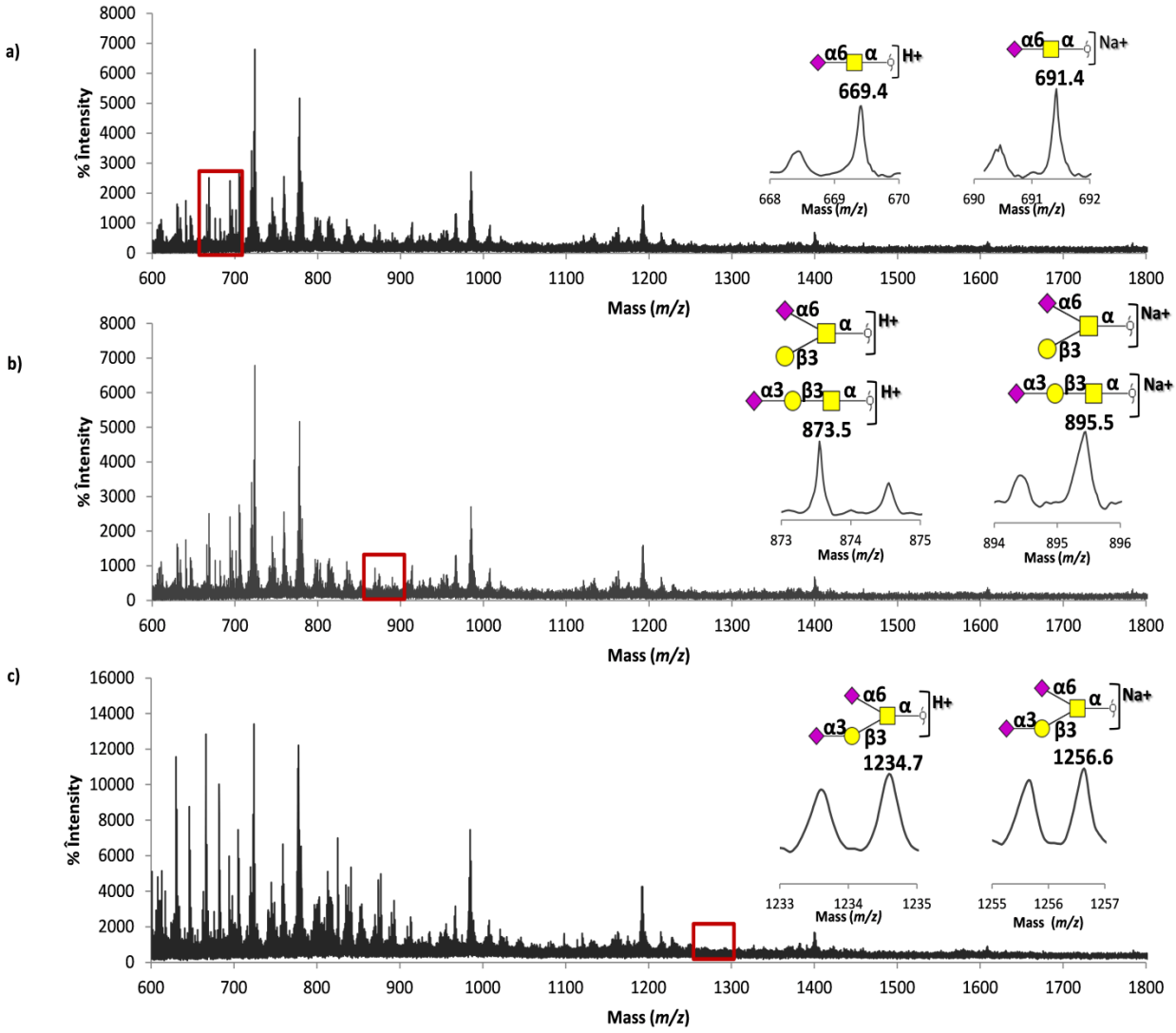


Figure 3. Nano-HPLC-MALDI-MS fractions exhibiting distinguishable signals for a) sialyl-Tn, b) sialyl-T (s-3-T and s-6-T), and c) disialyl-T $[M+H]^+$ and $[M+Na]^+$ ions. (○ - Gal; □ - GalNAc; ◆ - Neu5Ac)

Moreover, fluctuations that can go up to 20 ppm in relation to the expected monoisotopic masses were observed for a given ion within a chromatographic run. These deviations are most likely a consequence of lack of homogeneity in the crystallization of the matrix and changes in solvent composition during chromatography. To ensure the accuracy of the above assignments the glycans were digested with α-neuraminidase prior to permethylation. Such resulted in the loss of the chromatographic envelope for all ions and conditions tested (Figure 2c and d), thus confirming the existence of sialic acids. The subsequent MS^2 analyses, necessary to achieve inequivocous structural assignments, were governed by the results from the above described data mining technique. Namely, we selected LC fractions exhibiting proton and/or sodium adducts of sialyl-Tn, mono and di-sialyl-T with signal-to-noise ratios higher than 50 for MS^2 experiments. The characteristic product ion spectra for protonated and sodiated ions of the species under study have been summarized in Figure 4 and 5, respectively. They showed that, irrespectively of the type of adduct, there is a predominance of ions resulting from combination of X type cross-ring fragmentations with one or several B, C, Y and Z type fragmentations (nomenclature fragmentation according to Domon and Costello (1988) [37]). This array of fragmentations occurs at both the reducing and non-reducing terminals and results from the high degree of energy induced by CID and vacuum conditions [54, 55]. Despite these constrains, all shown spectra presented fragmentation patterns consistent with the ions under study. Of note, the product ion spectra of the protonated form of sialyl-T (Figure 4b), exhibits the ions at *m/z* 364.2, 408.2 and 713.4 resulting from combined fragmentations of Gal and sialic acids linked to a GalNAc residue, thus characteristic of s-6-T. However, no specific reporter ions have been found for s-3-T that has also been described in fetuin. Conversely, the sodiated counterpart (Figure 5) exhibits a low intensity ion at *m/z* 477.2 and another at *m/z* 733.3 consistent with sialic acids linked to Gal residues only present in the s-3-T isomer. No signals attributable solely to s-6-T could be found. These observations suggest that a careful evaluation of both protonated and sodiated adducts should be undertaken to achieve full structural assignments.

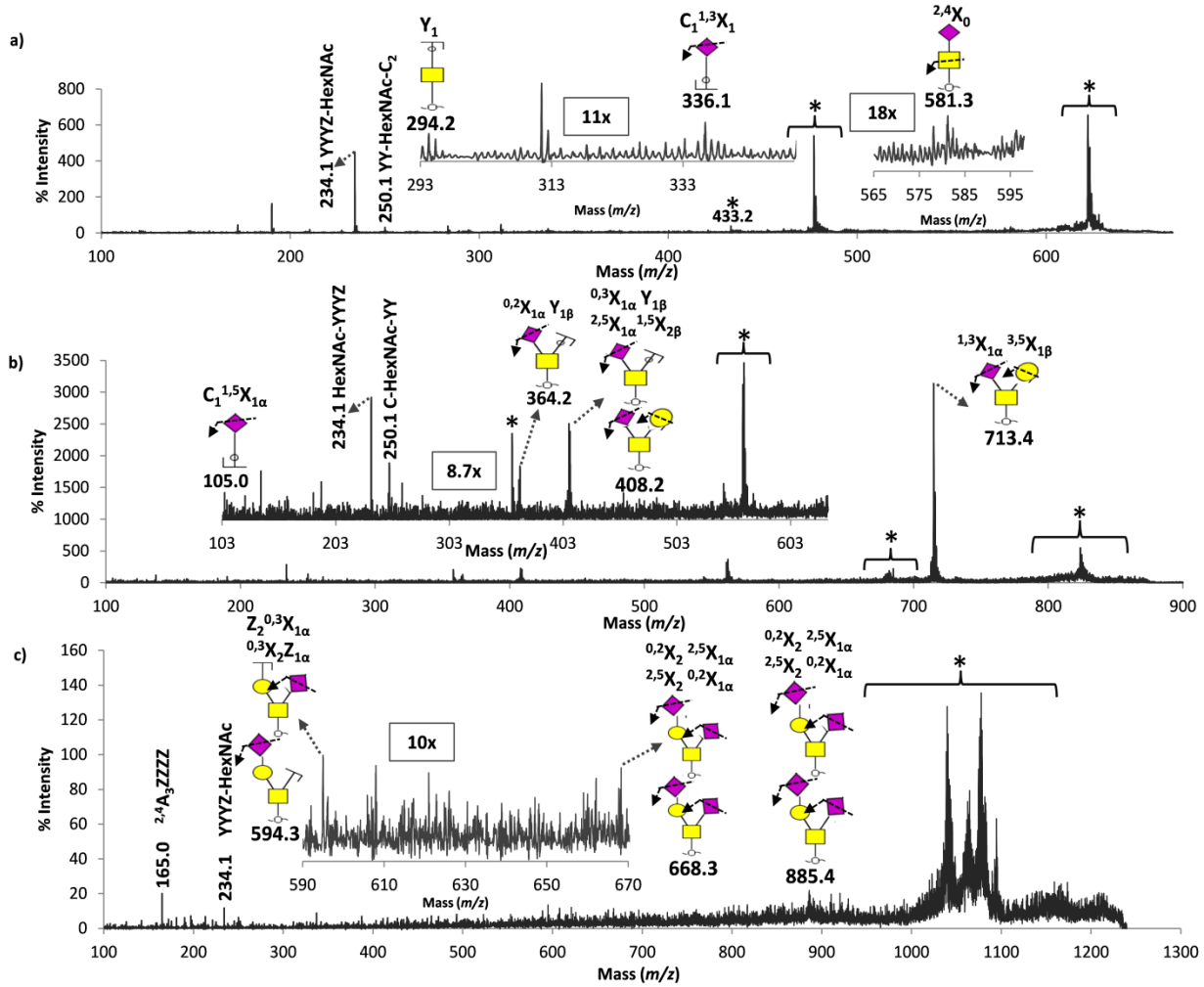


Figure 4. MALDI-TOF/TOF CID spectra of $[M+H]^+$ permethylated a) sialyl-Tn, b) sialyl-6-T, and c) disialyl-T exhibiting characteristic fragmentations according to the nomenclature introduced by Domon and Costello (1988) [37]. "*" signals resulting from the combination of cross-ring and glycosidic bound fragmentations occurring at both the reducing and non-reducing ends. (○ - Gal; □ - GalNAc; ◆ - Neu5Ac)

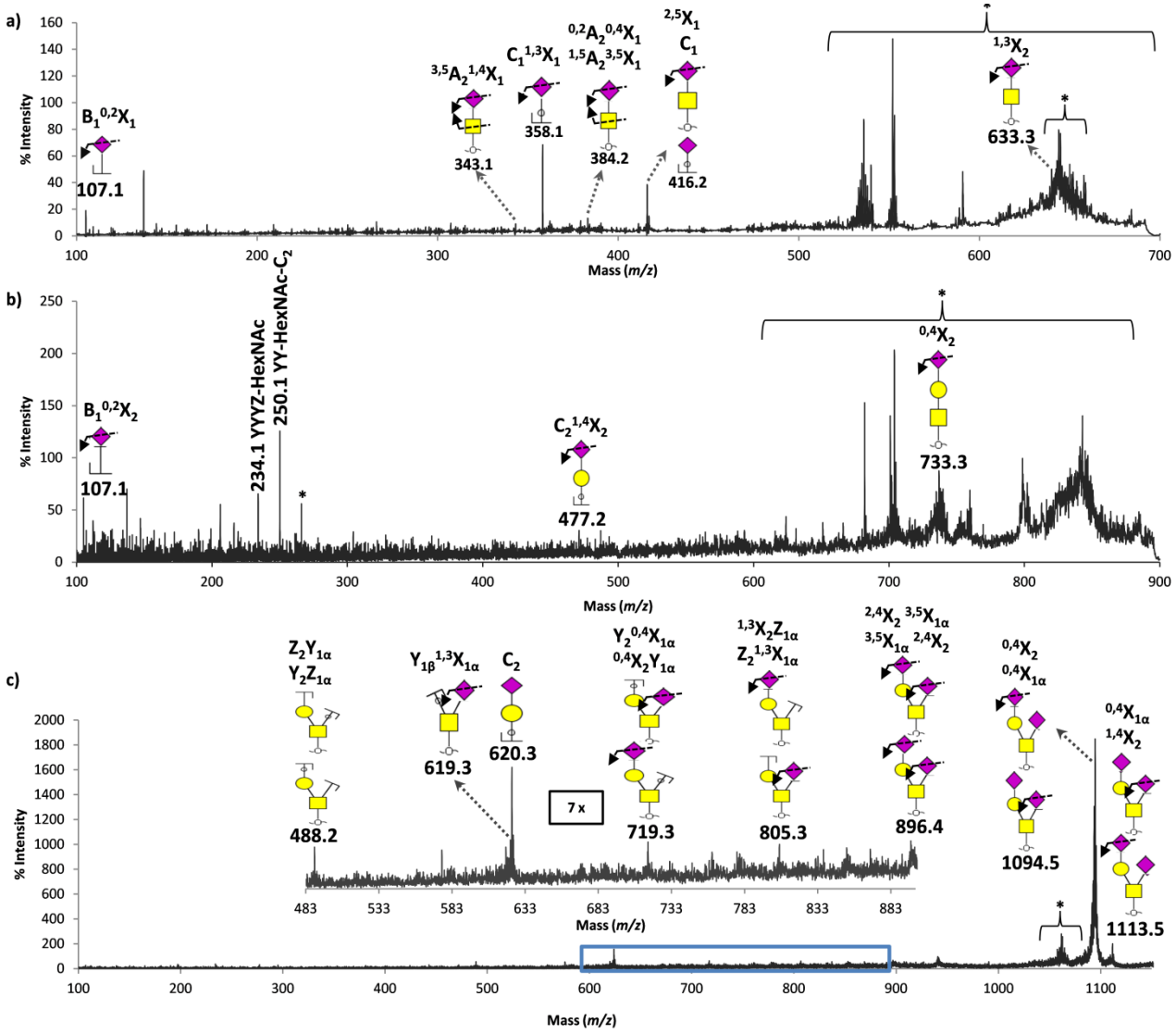


Figure 5. MALDI-TOF/TOF CID spectra of $[M+Na]^+$ permethylated a) sialyl-Tn, b) sialyl-T, and c) disialyl-T exhibiting characteristic fragmentations according to the nomenclature introduced by Domon and Costello (1988) [37]. "*" signals resulting from the combination of cross-ring and glycosidic bound fragmentations occurring at both the reducing and non-reducing ends. (○ - Gal; □ - GalNAc; ◆ - Neu5Ac)

In summary, it has been demonstrated that an hyphenated approach using in-gel de-O-glycosylation, permethylation and nano-HPLC-MALDI-TOF/TOF backed by comprehensive bioinformatics allows the detection of sialylated TF-related antigens from low *pico*mole to *femto*mole amounts of a glycoprotein of interest. Noteworthy, this technique also enabled the assignment of low molecular weight glycans found in zones of the spectra with strong background noise generated by acidic matrices such as DHB [56]. This was the case of the sialyl-Tn whose analysis by MALDI is being described, to our knowledge, for the first time. It should however be pointed out that, in those cases, a prior knowledge of the ions of interest is necessary, making this a target-driven approach.

- *Determination of sialylated TF-related antigens in bladder-cancer associated glycoproteins*

Despite the biological importance of sialyl-Tn in human malignancies [2, 13] and the ongoing efforts to develop cancer vaccines using this epitope [57], scarce information has been provided regarding its role in bladder tumors. Within the scope of our current research on this topic, we have performed immunohistochemistry staining against sialyl-Tn on serial FFPE rat urothelium sections exhibiting no-tumors and chemically-induced tumors. These studies showed no reactivity between the anti-sialyl Tn monoclonal antibody TKH2 and the healthy urothelium, demonstrating that this glycan is not expressed. Conversely, several tumors exhibited membrane and, to some extent cytoplasmatic immunostaining, consistent with the sialyl-Tn expression pattern [58, 59] (Figure 6a). We have further hypothesized that the glycoproteins carrying this modification are likely to be shed into the urine. Based on these observations, we have collected the urines of healthy rats and also those presenting sialyl-Tn positive tumors. The proteins on these urines were further purified by dialysis, combined in two pools (healthy and bladder tumor) of equal amounts, separated by SDS-PAGE, stained with Colloidal Coomassie Blue and blotted for sialyl-Tn. This highlighted a band at 25 kDa that was solely observable in the sialyl-Tn positive bladder tumors associated urines (Figure 6b). Moreover, this band showed considerable cross-reactivity to TKH2 antibody strongly suggesting the presence of sialyl-Tn (Figure 6b). The amount of material in this band obtained by comparing the intensities of staining with bovine serum albumin standards applied in the same gel was

roughly estimated to be 60 *pico*mol (1.5 µg). As described for fetuin, the band of interest was further isolated, the glycans released by chemically-assisted de-O-glycosylation, permethylated and analyzed by nano-HPLC-MALDI-TOF/TOF. Likewise, negative controls consisting of de-sialylated glycoprotein have also been analyzed. As shown by the plots in Figure 3c and d, the tumor-associated urinary glycoprotein exhibited a chromatographic envelope over the *m/z* values of both protonated and sialylated forms of sialyl-Tn (Figure 6c), that disappeared after treatment with the α-neuraminidase (Figure 3a and c). These assignments were further validated by MS^2 (spectrum similar to the one presented in Figure 4a) thus confirming blotting observations. This constitutes the first description of sialyl-Tn in urine in the context of bladder cancer. Studies are ongoing to determine the identity of the glycoproteins carrying this abnormal posttranslational modification and validate the association with bladder cancer. Noteworthy, nothing is known about the patterns of sialyl-Tn in human bladder tumors and its clinical value.

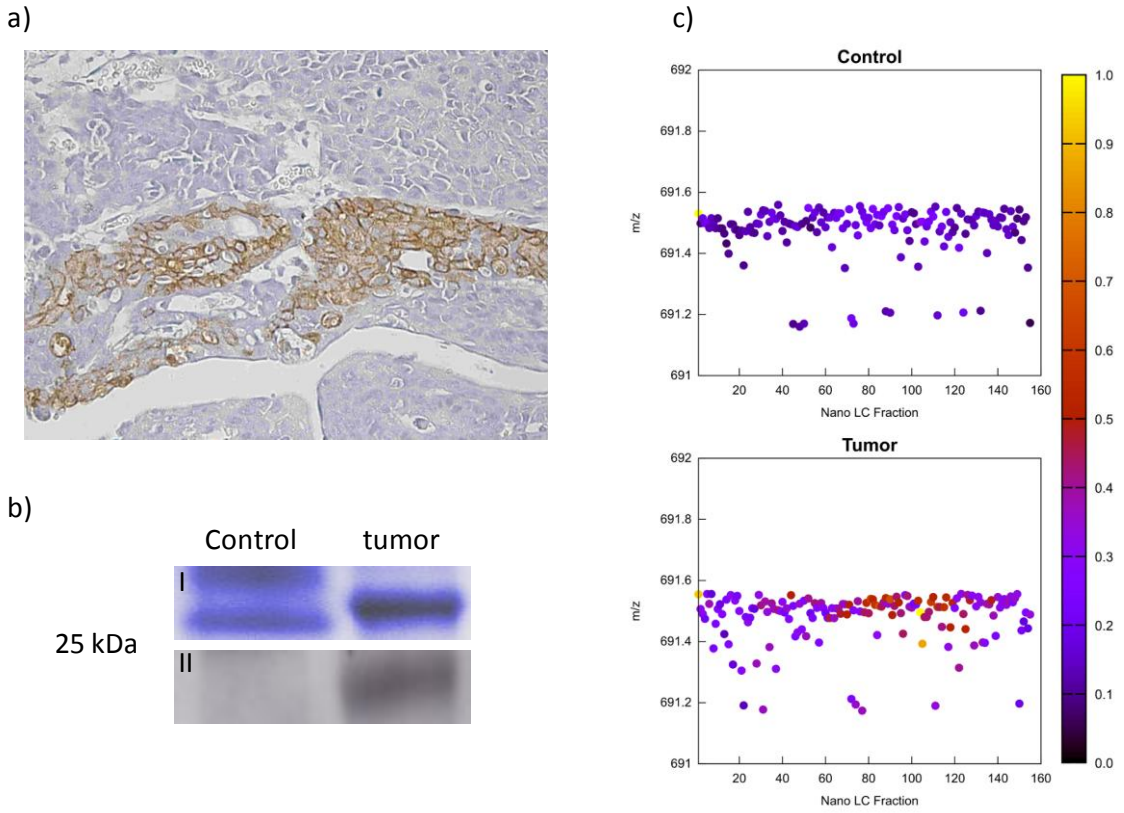


Figure 6. a) Rat urothelium with BBN-induced tumor showing intense immunostaining for sialyl-Tn; b) SDS-PAGE gel section from urine proteins showing a tumor-specific band at 25 kDa (I) immunoreactive with anti-sialyl Tn monoclonal antibody TKH2 (II); c) Plots of the sparse matrix (Fraction Number *vs m/z vs* nano-HPLC-MALDI-MS signal) for the peaks found within 669.4±0.2 Da, which includes the monoisotopic mass of protonated form of sialyl-Tn.

CONCLUSIONS

The assignment of sialylated TF-related antigens by MALDI poses a major analytical constrains that have been partially overcome by permethylation of the native glycans. Still, this derivatization is far from being a straightforward approach as often produces several undesirable byproducts. As a result it has been mostly applied when considerable amounts of glycoproteins are available. Only one study has suggested its application within the *pico*mole range, yet no attempts have been made to analyze glycans expressed in low abundance and isolated from gel spots [60]. Our study has now

demonstrated that this technique can be scaled down to glycoprotein amounts higher than 1.6 *pico*mol based on hyphenated approach involving permethylation and nano-HPLC-MALDI-TOF/TOF backed by comprehensive bioinformatics. Moreover, it has been applied successfully to glycoproteins isolated in gel spots which has been regarded as one of the main limitations of permethylation. This analytical approach will allow the validation of immune-based assignments from low amounts of starting material and, as such, of great interest those working in the field of glycobiology and proteomics.

REFERENCES

1. Reis, C.A., H. Osorio, L. Silva, C. Gomes, and L. David, *Alterations in glycosylation as biomarkers for cancer detection.* J Clin Pathol, 2010. **63**(4): p. 322-9.
2. Brockhausen, I., *Pathways of O-glycan biosynthesis in cancer cells.* Biochim Biophys Acta, 1999. **1473**(1): p. 67-95.
3. Yonezawa, S., T. Tachikawa, S. Shin, and E. Sato, *Sialosyl-Tn antigen. Its distribution in normal human tissues and expression in adenocarcinomas.* Am J Clin Pathol, 1992. **98**(2): p. 167-74.
4. Ikeda, Y., H. Kuwano, K. Baba, M. Ikebe, T. Matushima, Y. Adachi, M. Mori, and K. Sugimachi, *Expression of Sialyl-Tn antigens in normal squamous epithelium, dysplasia, and squamous cell carcinoma in the esophagus.* Cancer Res, 1993. **53**(7): p. 1706-8.
5. Itzkowitz, S.H., E.J. Bloom, W.A. Kokal, G. Modin, S. Hakomori, and Y.S. Kim, *Sialosyl-Tn - a Novel Mucin Antigen Associated with Prognosis in Colorectal-Cancer Patients.* Cancer, 1990. **66**(9): p. 1960-1966.
6. Kim, G.E., H.I. Bae, H.U. Park, S.F. Kuan, S.C. Crawley, J.J. Ho, and Y.S. Kim, *Aberrant expression of MUC5AC and MUC6 gastric mucins and sialyl Tn antigen in intraepithelial neoplasms of the pancreas.* Gastroenterology, 2002. **123**(4): p. 1052-60.
7. David, L., J.M. Nesland, H. Clausen, F. Carneiro, and M. Sobrinho-Simoes, *Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases.* APMIS Suppl, 1992. **27**: p. 162-72.
8. Inoue, M., H. Ogawa, O. Tanizawa, Y. Kobayashi, M. Tsujimoto, and T. Tsujimura, *Immunodetection of sialyl-Tn antigen in normal, hyperplastic and cancerous tissues of the uterine endometrium.* Virchows Arch A Pathol Anat Histopathol, 1991. **418**(2): p. 157-62.
9. Nishiyama, T., Y. Matsumoto, H. Watanabe, M. Fujiwara, and S. Sato, *Detection of Tn antigen with Vicia villosa agglutinin in urinary bladder cancer: its relevance to the patient's clinical course.* J Natl Cancer Inst, 1987. **78**(6): p. 1113-8.
10. Kobayashi, H., T. Terao, and Y. Kawashima, *Serum sialyl Tn as an independent predictor of poor prognosis in patients with epithelial ovarian cancer.* J Clin Oncol, 1992. **10**(1): p. 95-101.

11. Marcos, N.T., E.P. Bennett, J. Gomes, A. Magalhaes, C. Gomes, L. David, I. Dar, C. Jeanneau, S. DeFrees, D. Krustrup, L.K. Vogel, E.H. Kure, J. Burchell, J. Taylor-Papadimitriou, H. Clausen, U. Mandel, and C.A. Reis, *ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues.* Front Biosci (Elite Ed), 2011. **3**: p. 1443-55.
12. Ogata, S., P.J. Maimonis, and S.H. Itzkowitz, *Mucins bearing the cancer-associated sialosyl-Tn antigen mediate inhibition of natural killer cell cytotoxicity.* Cancer Res, 1992. **52**(17): p. 4741-6.
13. Pinho, S., N.T. Marcos, B. Ferreira, A.S. Carvalho, M.J. Oliveira, F. Santos-Silva, A. Harduin-Lepers, and C.A. Reis, *Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells.* Cancer Lett, 2007. **249**(2): p. 157-70.
14. Liotta, L.A. and E.C. Kohn, *The microenvironment of the tumour-host interface.* Nature, 2001. **411**(6835): p. 375-9.
15. Werther, J.L., S. Rivera-MacMurray, H. Bruckner, M. Tatematsu, and S.H. Itzkowitz, *Mucin-associated sialosyl-Tn antigen expression in gastric cancer correlates with an adverse outcome.* Br J Cancer, 1994. **69**(3): p. 613-6.
16. Nakagoe, T., T. Sawai, T. Tsuji, M.A. Jibiki, A. Nanashima, H. Yamaguchi, T. Yasutake, H. Ayabe, K. Arisawa, and H. Ishikawa, *Predictive factors for preoperative serum levels of sialy Lewis(x), sialyl Lewis(a) and sialyl Tn antigens in gastric cancer patients.* Anticancer Res, 2002. **22**(1A): p. 451-8.
17. Itzkowitz, S.H., E.J. Bloom, T.S. Lau, and Y.S. Kim, *Mucin associated Tn and sialosyl-Tn antigen expression in colorectal polyps.* Gut, 1992. **33**(4): p. 518-23.
18. Holmberg, L.A., D.V. Oparin, T. Gooley, K. Lilleby, W. Bensinger, M.A. Reddish, G.D. MacLean, B.M. Longenecker, and B.M. Sandmaier, *Clinical outcome of breast and ovarian cancer patients treated with high-dose chemotherapy, autologous stem cell rescue and THERATOPE STn-KLH cancer vaccine.* Bone Marrow Transplant, 2000. **25**(12): p. 1233-41.
19. Rajpert-De Meyts, E., S.N. Poll, I. Goukasian, C. Jeanneau, A.S. Herlihy, E.P. Bennett, N.E. Skakkebaek, H. Clausen, A. Giwercman, and U. Mandel, *Changes in the profile of simple mucin-type O-glycans and polypeptide GalNAc-transferases in human testis and testicular neoplasms are associated with germ cell maturation and tumour differentiation.* Virchows Arch, 2007. **451**(4): p. 805-14.
20. Videira, P.A., M. Correia, N. Malagolini, H.J. Crespo, D. Ligeiro, F.M. Calais, H. Trindade, and F. Dall'Olio, *ST3Gal.I sialyltransferase relevance in bladder cancer tissues and cell lines.* BMC Cancer, 2009. **9**: p. 357.
21. Beum, P.V., J. Singh, M. Burdick, M.A. Hollingsworth, and P.W. Cheng, *Expression of core 2 beta-1,6-N-acetylglucosaminyltransferase in a human pancreatic cancer cell line results in altered expression of MUC1 tumor-associated epitopes.* J Biol Chem, 1999. **274**(35): p. 24641-8.
22. Feizi, T., *Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens.* Nature, 1985. **314**(6006): p. 53-7.
23. Pinto, R., A.S. Carvalho, T. Conze, A. Magalhaes, G. Picco, J.M. Burchell, J. Taylor-Papadimitriou, C.A. Reis, R. Almeida, U. Mandel, H. Clausen, O. Soderberg, and L. David, *Identification of new cancer biomarkers based on aberrant mucin glycoforms by in situ Proximity Ligation.* J Cell Mol Med, 2011.

24. Kyselova, Z., Y. Mechref, P. Kang, J.A. Goetz, L.E. Dobrolecki, G.W. Sledge, L. Schnaper, R.J. Hickey, L.H. Malkas, and M.V. Novotny, *Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles.* Clin Chem, 2008. **54**(7): p. 1166-75.
25. de Leoz, M.L., L.J. Young, H.J. An, S.R. Kronewitter, J. Kim, S. Miyamoto, A.D. Borowsky, H.K. Chew, and C.B. Lebrilla, *High-mannose glycans are elevated during breast cancer progression.* Mol Cell Proteomics, 2011. **10**(1): p. M110 002717.
26. Tang, Z., R.S. Varghese, S. Bekesova, C.A. Loffredo, M.A. Hamid, Z. Kyselova, Y. Mechref, M.V. Novotny, R. Goldman, and H.W. Ressom, *Identification of N-glycan serum markers associated with hepatocellular carcinoma from mass spectrometry data.* J Proteome Res, 2010. **9**(1): p. 104-12.
27. Williams, T.I., D.A. Saggese, K.L. Toups, J.L. Frahm, H.J. An, B. Li, C.B. Lebrilla, and D.C. Muddiman, *Investigations with O-linked protein glycosylations by matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry.* J Mass Spectrom, 2008. **43**(9): p. 1215-23.
28. Zaia, J., *Mass spectrometry and glycomics.* OMICS, 2010. **14**(4): p. 401-18.
29. Moyer, S.C. and R.J. Cotter, *Atmospheric pressure MALDI.* Anal Chem, 2002. **74**(17): p. 468A-476A.
30. O'Connor, P.B., E. Mirgorodskaya, and C.E. Costello, *High pressure matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for minimization of ganglioside fragmentation.* J Am Soc Mass Spectrom, 2002. **13**(4): p. 402-7.
31. Demelbauer, U.M., M. Zehl, A. Plematl, G. Allmaier, and A. Rizzi, *Determination of glycopeptide structures by multistage mass spectrometry with low-energy collision-induced dissociation: comparison of electrospray ionization quadrupole ion trap and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron time-of-flight approaches.* Rapid Commun Mass Spectrom, 2004. **18**(14): p. 1575-82.
32. Dell, A., *Preparation and desorption mass spectrometry of permethyl and peracetyl derivatives of oligosaccharides.* Methods Enzymol, 1990. **193**: p. 647-60.
33. Ressom, H.W., R.S. Varghese, L. Goldman, C.A. Loffredo, M. Abdel-Hamid, Z. Kyselova, Y. Mechref, M. Novotny, and R. Goldman, *Analysis of MALDI-TOF mass spectrometry data for detection of glycan biomarkers.* Pac Symp Biocomput, 2008: p. 216-27.
34. Viseux, N., E. de Hoffmann, and B. Domon, *Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry.* Anal Chem, 1998. **70**(23): p. 4951-9.
35. Viseux, N., E. de Hoffmann, and B. Domon, *Structural analysis of permethylated oligosaccharides by electrospray tandem mass spectrometry.* Anal Chem, 1997. **69**(16): p. 3193-8.
36. Zaia, J., *Mass spectrometry of oligosaccharides.* Mass Spectrom Rev, 2004. **23**(3): p. 161-227.
37. Domon, B. and C.E. Costello, *A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-Ms Ms Spectra of Glycoconjugates.* Glycoconjugate Journal, 1988. **5**(4): p. 397-409.
38. Mechref, Y., P. Kang, and M.V. Novotny, *Differentiating structural isomers of sialylated glycans by matrix-assisted laser desorption/ionization time-of-*

flight/time-of-flight tandem mass spectrometry. Rapid Commun Mass Spectrom, 2006. **20**(8): p. 1381-9.
39. Yu, S.Y., S.W. Wu, and K.H. Khoo, *Distinctive characteristics of MALDI-Q/TOF and TOF/TOF tandem mass spectrometry for sequencing of permethylated complex type N-glycans.* Glycoconj J, 2006. **23**(5-6): p. 355-69.
40. Pang, P.C., B. Tissot, E.Z. Drobnis, P. Sutovsky, H.R. Morris, G.F. Clark, and A. Dell, *Expression of bisecting type and Lewisx/Lewisy terminated N-glycans on human sperm.* J Biol Chem, 2007. **282**(50): p. 36593-602.
41. Wuhrer, M., C.A. Koeleman, C.H. Hokke, and A.M. Deelder, *Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments.* Rapid Commun Mass Spectrom, 2006. **20**(11): p. 1747-54.
42. Ferreira, J.A., A.L. Daniel-da-Silva, R.M. Alves, D. Duarte, I. Vieira, L.L. Santos, R. Vitorino, and F. Amado, *Synthesis and optimization of lectin functionalized nanoprobes for the selective recovery of glycoproteins from human body fluids.* Anal Chem, 2011. **83**(18): p. 7035-43.
43. Oliveira, P.A., F. Adega, C.A. Palmeira, R.M. Chaves, A.A. Colaco, H. Guedes-Pinto, P.L. De la Cruz, and C.A. Lopes, *DNA study of bladder papillary tumours chemically induced by N-butyl-N-(4-hydroxybutyl) nitrosamine in Fisher rats.* Int J Exp Pathol, 2007. **88**(1): p. 39-46.
44. Halim, A., J. Nilsson, U. Ruetschi, C. Hesse, and G. Larson, *Human Urinary Glycoproteomics; Attachment Site Specific Analysis of N- and O-Linked Glycosylations by CID and ECD.* Mol Cell Proteomics, 2012. **11**(4): p. M111 013649.
45. Ciucanu, I. and F. Kerek, *A Simple and Rapid Method for the Permethylation of Carbohydrates.* Carbohydrate Research, 1984. **131**(2): p. 209-217.
46. Ceroni, A., K. Maass, H. Geyer, R. Geyer, A. Dell, and S.M. Haslam, *GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of Glycans.* Journal of Proteome Research, 2008. **7**(4): p. 1650-1659.
47. Wada, Y., P. Azadi, C.E. Costello, A. Dell, R.A. Dwek, H. Geyer, R. Geyer, K. Kakehi, N.G. Karlsson, K. Kato, N. Kawasaki, K.H. Khoo, S. Kim, A. Kondo, E. Lattova, Y. Mechref, E. Miyoshi, K. Nakamura, H. Narimatsu, M.V. Novotny, N.H. Packer, H. Perreault, J. Peter-Katalinic, G. Pohlentz, V.N. Reinhold, P.M. Rudd, A. Suzuki, and N. Taniguchi, *Comparison of the methods for profiling glycoprotein glycans - HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study.* Glycobiology, 2007. **17**(4): p. 411-422.
48. Merry, A.H., D.C. Neville, L. Royle, B. Matthews, D.J. Harvey, R.A. Dwek, and P.M. Rudd, *Recovery of intact 2-aminobenzamide-labeled O-glycans released from glycoproteins by hydrazinolysis.* Anal Biochem, 2002. **304**(1): p. 91-9.
49. Royle, L., T.S. Mattu, E. Hart, J.I. Langridge, A.H. Merry, N. Murphy, D.J. Harvey, R.A. Dwek, and P.M. Rudd, *An analytical and structural database provides a strategy for sequencing O-glycans from microgram quantities of glycoproteins.* Anal Biochem, 2002. **304**(1): p. 70-90.
50. Nwosu, C.C., R.R. Seipert, J.S. Strum, S.S. Hua, H.J. An, A.M. Zivkovic, B.J. German, and C.B. Lebrilla, *Simultaneous and extensive site-specific N- and O-glycosylation analysis in protein mixtures.* J Proteome Res, 2011. **10**(5): p. 2612-24.

51. Iskratsch, T., A. Braun, K. Paschinger, and I.B.H. Wilson, *Specificity analysis of lectins and antibodies using remodeled glycoproteins.* Analytical Biochemistry, 2009. **386**(2): p. 133-146.
52. Green, E.D., G. Adelt, J.U. Baenziger, S. Wilson, and H. Van Halbeek, *The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy.* J Biol Chem, 1988. **263**(34): p. 18253-68.
53. Cancilla, M.T., S.G. Penn, J.A. Carroll, and C.B. Lebrilla, *Coordination of alkali metals to oligosaccharides dictates fragmentation behavior in matrix assisted laser desorption ionization Fourier transform mass spectrometry.* Journal of the American Chemical Society, 1996. **118**(28): p. 6736-6745.
54. Stephens, E., S.L. Maslen, L.G. Green, and D.H. Williams, *Fragmentation characteristics of neutral N-linked glycans using a MALDI-TOF/TOF tandem mass spectrometer.* Anal Chem, 2004. **76**(8): p. 2343-54.
55. Mechref, Y., M.V. Novotny, and C. Krishnan, *Structural characterization of oligosaccharides using MALDI-TOF/TOF tandem mass spectrometry.* Anal Chem, 2003. **75**(18): p. 4895-903.
56. Hashir, M.A., G. Stecher, and G.K. Bonn, *Identification of low molecular weight carbohydrates employing new binary mixtures for matrix-assisted laser desorption/ionisation mass spectrometry.* Rapid Communications in Mass Spectrometry, 2008. **22**(14): p. 2185-2194.
57. Heimburg-Molinaro, J., M. Lum, G. Vijay, M. Jain, A. Almogren, and K. Rittenhouse-Olson, *Cancer vaccines and carbohydrate epitopes.* Vaccine, 2011. **29**(48): p. 8802-26.
58. Soares, R., A. Marinho, and F. Schmitt, *Expression of Sialyl-Tn in breast cancer - Correlation with prognostic parameters.* Pathology Research and Practice, 1996. **192**(12): p. 1181-1186.
59. Yoshida, A., M. Sotozono, T. Nakatou, Y. Okada, and T. Tsuji, *Different expression of Tn and sialyl-Tn antigens between normal and diseased human gastric epithelial cells.* Acta Medica Okayama, 1998. **52**(4): p. 197-204.
60. Mechref, Y., P. Kang, and M.V. Novotny, *Solid-phase permethylation for glycomic analysis.* Methods Mol Biol, 2009. **534**: p. 53-64.

Chapter IV. Detection of sialyl-Tn in plasminogen from serum of patients with gastric cancer precursor lesions by in gel deglycosylation and nanoLC-MALDI-TOF/TOF

ABSTRACT

Gastric cancer is a global health burden and the second cause of cancer related mortality worldwide. It is generally associated to the development of gastritis, that can lead to chronic atrophic gastritis, and intestinal metaplasia which are regarded as percursors of gastric adenocarcinoma. Currently, the diagnostic for these primary lesions relies almost exclusively in endoscopy followed by biopsy, which is both invasive and costly to apply in a routine basis. Therefore, biomarkers which can assist the non invasive screening for silent gastric pathologies are highly needed. In this context, our current glycoproteomic-based studies by western blot have shown an association between serum plasminogen carrying the tumor-associated carbohydrate sialyl-Tn antigen and the above mentioned percursor stages.

To confirm these findings, plasminogen was enriched from serum of patients with gastritis and intestinal metaplasia by lysine-sepharose affinity chromatography and further isolated by gel electrophoresis. The presence of sialyl-Tn was then accessed by nanoLC-MALDI-TOF/TOF using the methodology described in chapter III of this dissertation. This approach allowed the validation of western blot assignments thereby setting the structural basis for a novel biomarker for the early detection of gastric cancer.

Keywords: sialylated TF-related antigens, gastritis, intestinal metaplasia, sialyl-Tn, plasminogen, permethylation

INTRODUCTION

Gastric cancer is a global health burden and the second cause of cancer related mortality worldwide. The development of gastric cancer is associated with a long carcinogenesis pathway generally initiated by the Gram-negative bacterium *Helicobacter pylori* [2]. Infection may result in the development of gastritis, that can lead to chronic atrophic gastritis, and intestinal metaplasia which are regarded as precursors of gastric adenocarcinoma [3]. Currently, the diagnostic for these precursor lesions relies almost exclusively in endoscopy followed by biopsy, which is both invasive and costly to apply in a routine basis [4]. Therefore, biomarkers which can assist the non invasive screening for silent gastric pathologies are highly needed.

Glycosylation is a common post-translational modification of proteins with more than half of eukaryotic proteins thought to be glycosylated [5]. The pattern of protein glycosylation is cell and tissue specific and closely reflects the physiological status of the cell [6]. Thus, changes in glycan expression are being observed for several pathological conditions and in particular in gastric lesions [7]. Those include the aberrant expression of simple mucin type carbohydrate antigens, namely T (Galβ1–3GalNAcα-O-Ser/Thr), Tn (GalNAcα-O-Ser/Thr), and sialyl-Tn (Neu5Acalpha2-6GalNAcα-O-Ser/Thr) [8-11]. An overexpression of sialylated Lewis antigens [12] and the decreased expression of terminal αGlcNAc O-glycans [10, 13, 14] has also been reported. As a result, some of the most promising biomarkers for arise from alterations in the glycosylation pathways accompanying oncogenic transformations. Protein carrying these alterations can ultimately be released into the blood stream and constitute a valuable biomarkers [15]. Indeed, the majority of the serological cancer biomarkers known to date are either glycans at cell surface of the tumor (attached to glycoproteins or glycolipids) or in secreted glycoproteins. Among the most proheminent serological assays are the ones targeting carbohydrate antigens sialyl Lewis a (CA19-9) and sialyl-Tn (CA72-4) or mucin glycoproteins such as MUC1 (CA15-3) and MUC16 (CA125) [7].

In this context, our current glycoproteomic-based studies by western blot have shown an association between serum plasminogen carrying the tumor associated carbohydrate sialyl-Tn (Neu5Acα2-6GalNAc-*O*-Ser/Thr) antigen gastric cancer precursor lesions. In line with these observations sialyl-Tn has been previously observed in pre-malignant lesions of the gastrointestinal tract, namely in intestinal metaplasia [16].

Acknowledging the biomarker value of these observations, the present work is devoted to the validate the western blot assignments.

As such, we have isolated plasminogen from serum samples associated with gastritis and the two types of intestinal metaplasia (complete and incomplete). Using the analytical approach described in chapter III it was possible to detect by MALDI-TOF/TOF sialyl-Tn in the plasminogen associated with gastric pathologies, thus validating western blot data.

EXPERIMENTAL METHODS

Serum collection and enrichment of Plasminogen

The plasminogen was isolated from serum of several patients with gastritis, intestinal metaplasia **(**complete and incomplete**)** as well as healthy donors using lysine sepharose affinity chromatography as described by Deutsch, and Mertz (1970) [17].The beads of lysine sepharose (200µL) were washed 5 times with 2-3 bed volumes of binding buffer (phosphate 50mM; pH 7.5) and incubated with 150µL serum and 400µL of binding buffer overnight to a 4ºC. In the next day the beads were washed 5 times with binding buffer and eluted with 200µL elution buffer (phosphate 50mM; pH 7.5, 0.2M of aminocaproic acid) 3 times. The protein content in each sample was determined using RC-DC protein quantification Kit (BioRAD). The recovered proteins were then separated by SDS-PAGE (9% Mini-PROTEAN TGX Precast Gels (BIORAD)) and stained with Coomassie Blu**e**. The plasminogen bands were removed from SDS-PAGE gels and the protein identity was confirmed by MALDI MS. Equivalent amounts of plasminogen previously digested with neuraminidase from *Clostridium perfringens* (Sigma-Aldrich; Karlsruhe, Germany) were used as a control.

Protein selection and in-gel tryptic digestion.

One dimension gels were stained with Coomassie Blue (Bio-Safe Coomassie from Bio-Rad, CA) overnight and images were acquired with a Gel Doc XR system (Bio-Rad, CA). The spots corresponding to plasminogen were removed from the gel, destained, and

submitted to in-gel tryptic digestion and the peptides extracted according to the trypsin's manufacturer instructions (Promega, USA).

Protein identification by MALDI-TOF/TOF

The tryptic digests were separated with a C18 Pepmap (Dionex) column on an Ultimate 3000 (Dionex/LC Packings, Sunnyvale, CA) nano-HPLC and fractions were collected with a Probot (Dionex/LC Packings, Sunnyvale, CA) directly onto a matrix assisted laser desorption ionization (MALDI) plate. The MALDI-TOF/TOF (time-of-flight) mass spectrometry (MS) analysis was performed on a 4800 MALDI-TOF/TOF Analyzer (Applied Biosystems, Foster City, CA). The MS and MS/MS spectra acquired were processed and analyzed by the Global Protein Server Workstation (Applied Biosystems). Protein identification was achieved with a search performed against the Swiss-Prot protein database (March 2009, 428 650 entries) for Homo sapiens using the Mascot search engine (Version 2.1.04, Matrix Science, UK). The search included peaks with a signal-to-noise ratio greater than 10 and allowed for up to two missed trypsin cleavage sites. To be considered a match a confidence interval, calculated by the AB SCIEX GPS Explorer software, of at least 99% was required.

In gel de-O-glycosylation and permethylation and nanoLC-MALDI-TOF/TOF

The band corresponding to plasminogen was removed from SDS-PAGE gels and screened for sialyl-Tn by in-gel de-O-glycosylation, permethylation and nanoLC-MALDI-TOF/TOF using the methodology described in chapter III.

RESULTS AND DISCUSSION

Western blot screening for tumor-associated simple mucin type O-glycans in serum proteins as resulted in the identification of plasminogen carrying sialyl-Tn. Of particular interest, these abnormal O-glycosylated form of plasminogen has been mostly observed in gastritis and, in particular, intestinal metaplasia (complete and incomplete). Moreover, no expression was found among healthy individuals, reinforcing the putative value of this observation in the early detection of gastric cancer precursor lesions. To confirm blot-based assignment we have isolated plasminogen from serum samples associated with gastritis and the two types of intestinal metaplasia (complete and incomplete) as well as healthy donors. The presence of sialyl-Tn was then accessed using the analytical approach developed in chapter III comprising in gel reductive β-elimination, permethylation and nanoLC-MALDI-TOF/TOF.

The glycoprotein of interest was primarily enriched with lysine-sepharose from serum of patients with gastritis and intestinal metaplasia as well as from healthy donors. The lysine-sepharose is a resin with its amino and carboxyl groups free to interact with the negative charges of glutamic acid residues as well as Neu5Ac presents on plasminogen. The samples enriched in plasminogen were then separated in a gel electrophoresis SDS-PAGE 9% (Figure 1a). Then the spots at 100 kDa corresponding to plasminogen, as confirmed by protein identification by MALDI-TOF/TOF, were recovered from the gel and subjected to in-gel de-O-glycosylation by reductive β-elimination (Figure 1a). The resulting oligosaccharides were filtered for further permethylation with traces of water, according to Ciucanu and Kerek (1984) [18]. The permethylated glycans were separated by nano-HPLC using a C18 reverse phase column and analyzed by positive mode MALDI-TOF/TOF.

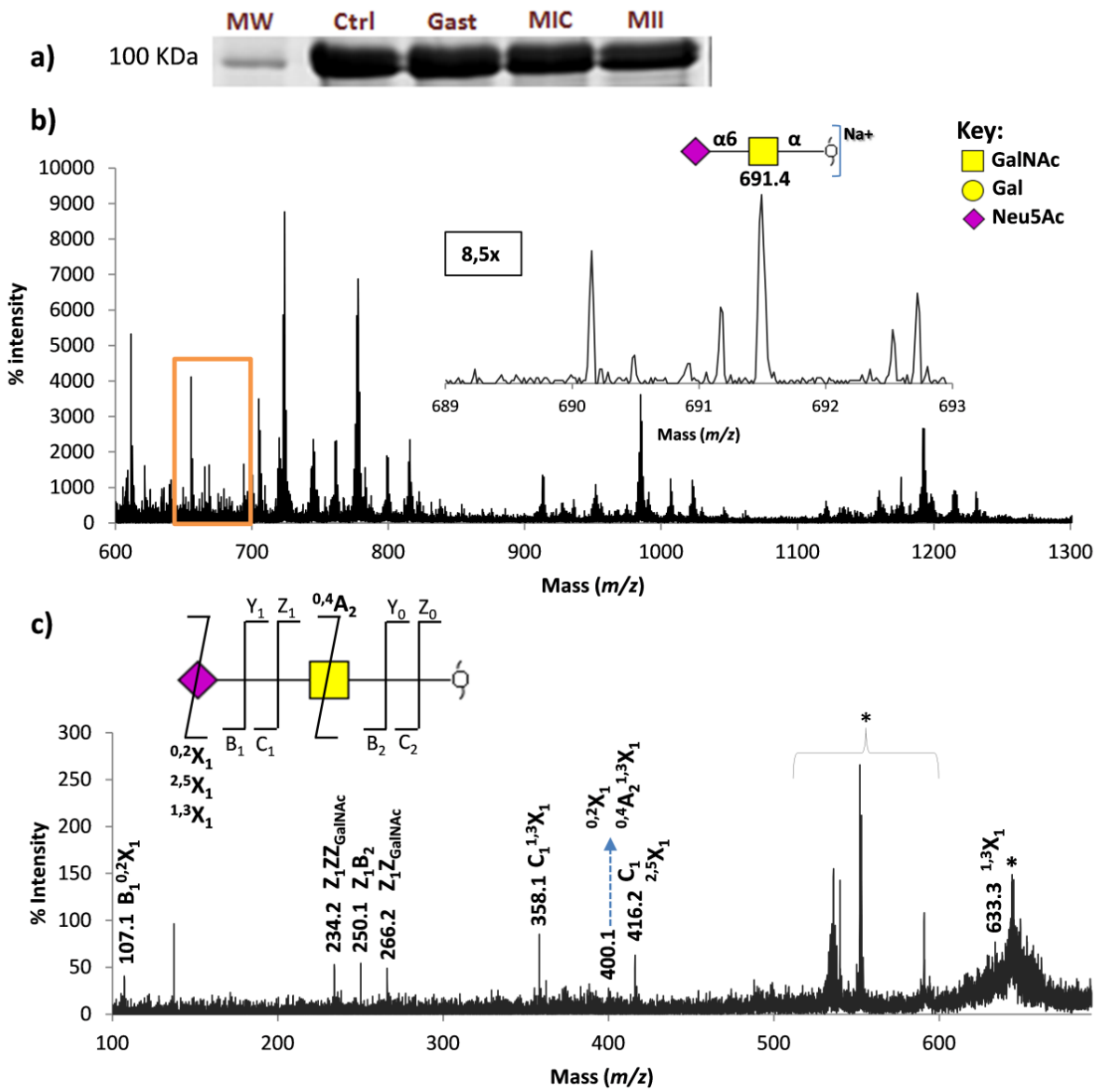


Figure 1. a) SDS-PAGE gel showing plasminogen from serum of healthy donors (Ctrl) and patients diagnosed with gastritis (Gast), complete (MIC) and incomplete (MII) intestinal metaplasia isolated by lysine-sepharose affinity chromatography and b) nano-HPLC-MALDI-MS fraction of permethylated sialyl-Tn from plasminogen of patients with incomplete metaplasia intestinal exhibiting $[M+Na]^+$ and c) MALDI-TOF/TOF spectra of $[M+Na]^+$ permethylated sialyl-Tn with unique fragmentations according to the nomenclature of Domon and Costello (1988) [1]. "*" signals resulting from the combination of cross-ring and glycosidic bound fragmentations occurring at both the reducing and non reducing ends.

As mentioned in the chapter III, a bioinformatics survey was necessary for the selection of the chromatographic runs corresponding to sialylated TF-related antigens. This allowed the distinction of an ion at *m/z* 691.4 corresponding to the sodium adduct of sialyl-Tn antigen (Figure 1b) in the plasminogen isolated from gastritis and intestinal metaplasia. The assignment was further reinforced by the disappearance of the signal upon treatment of plasminogen with a α-neuraminidase. Moreover, the product ion spectrum of the ion at *m/z* 691.4 exhibited glycosidic bond cleavages B, C and Z, and cross-ring fragmentations A and X (nomenclature according to Domon and Costello (1988) [1]) characteristic of sialyl-Tn (Figure 1C), therefore confirming the presence of sialyl-Tn in plasminogen. In agreement with previous reports concerning the O-glycosylation of plasminogen [19-21], low abundant ions belonging to mono (*m/z* 873.5 and 895.5) and disialylated (*m/z* 1234.7 and 1256.6) -TF antigens have also been detected.

CONCLUSION

The sialyl-Tn is considered serum biomarker for advanced gastric cancer [22, 23] and is frequently associated with malignant phenotypes, aggressiveness, invasion, metastasization, and poor prognosis [24, 25]. This epitope has been further explored in cancer vaccines and as target for finding new tumor associated glycoproteins. Our western blot-based glycoproteomics studies have now unveiled that plasminogen forms carrying sialyl-Tn could be found in the serum of patients with pre-cancerous gastric lesions. In particular, cases with both incomplete and complete intestinal metaplasia have shown considerable expression of this aberrantly O-glycosylated form of plasminogen. This work has now confirmed blot assignments by mass spectrometry setting the structural basis for further studies directed to access the clinical value of this observation in the early diagnostic of gastric cancer.

REFERENCES

1. Domon, B. and C.E. Costello, *A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-Ms Ms Spectra of Glycoconjugates.* Glycoconjugate Journal, 1988. **5**(4): p. 397-409.
2. Atherton, J.C., *The pathogenesis of Helicobacter pylori-induced gastro-duodenal diseases.* Annu Rev Pathol, 2006. **1**: p. 63-96.
3. Correa, P. and J. Houghton, *Carcinogenesis of Helicobacter pylori.* Gastroenterology, 2007. **133**(2): p. 659-72.
4. Sabbi, T., *Short review about Helicobacter pylori infection in pediatric age: epidemiological and clinical findings, diagnosis, therapy and role of probiotics.* Pediatr Med Chir, 2011. **33**(5-6): p. 221-6.
5. Apweiler, R., H. Hermjakob, and N. Sharon, *On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database.* Biochim Biophys Acta, 1999. **1473**(1): p. 4-8.
6. Brockhausen, I., *Pathways of O-glycan biosynthesis in cancer cells.* Biochim Biophys Acta, 1999. **1473**(1): p. 67-95.
7. Reis, C.A., H. Osorio, L. Silva, C. Gomes, and L. David, *Alterations in glycosylation as biomarkers for cancer detection.* J Clin Pathol, 2010. **63**(4): p. 322-9.
8. Conze, T., A.S. Carvalho, U. Landegren, R. Almeida, C.A. Reis, L. David, and O. Soderberg, *MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas.* Glycobiology, 2010. **20**(2): p. 199-206.
9. David, L., J.M. Nesland, H. Clausen, F. Carneiro, and M. Sobrinho-Simoes, *Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases.* APMIS Suppl, 1992. **27**: p. 162-72.
10. Ferreira, B., N.T. Marcos, L. David, J. Nakayama, and C.A. Reis, *Terminal alpha1,4-linked N-acetylglucosamine in Helicobacter pylori-associated intestinal metaplasia of the human stomach and gastric carcinoma cell lines.* J Histochem Cytochem, 2006. **54**(5): p. 585-91.
11. Marcos, N.T., E.P. Bennett, J. Gomes, A. Magalhaes, C. Gomes, L. David, I. Dar, C. Jeanneau, S. DeFrees, D. Krustrup, L.K. Vogel, E.H. Kure, J. Burchell, J. Taylor-Papadimitriou, H. Clausen, U. Mandel, and C.A. Reis, *ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues.* Front Biosci (Elite Ed), 2011. **3**: p. 1443-55.
12. Amado, M., F. Carneiro, M. Seixas, H. Clausen, and M. Sobrinho-Simoes, *Dimeric sialyl-Le(x) expression in gastric carcinoma correlates with venous invasion and poor outcome.* Gastroenterology, 1998. **114**(3): p. 462-70.
13. Karasawa, F., A. Shiota, Y. Goso, M. Kobayashi, Y. Sato, J. Masumoto, M. Fujiwara, S. Yokosawa, T. Muraki, S. Miyagawa, M. Ueda, M.N. Fukuda, M. Fukuda, K. Ishihara, and J. Nakayama, *Essential role of gastric gland mucin in preventing gastric cancer in mice.* J Clin Invest, 2012. **122**(3): p. 923-34.
14. Zhang, M.X., J. Nakayama, E. Hidaka, S. Kubota, J. Yan, H. Ota, and M. Fukuda, *Immunohistochemical demonstration of alpha1,4-N-acetylglucosaminyltransferase that forms GlcNAcalpha1,4Galbeta residues in human gastrointestinal mucosa.* J Histochem Cytochem, 2001. **49**(5): p. 587-96.

15. Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.
16. Karlen, P., E. Young, O. Brostrom, R. Lofberg, B. Tribukait, K. Ost, C. Bodian, and S. Itzkowitz, *Sialyl-Tn antigen as a marker of colon cancer risk in ulcerative colitis: relation to dysplasia and DNA aneuploidy.* Gastroenterology, 1998. **115**(6): p. 1395-404.
17. Deutsch, D.G. and E.T. Mertz, *Plasminogen: purification from human plasma by affinity chromatography.* Science, 1970. **170**(3962): p. 1095-6.
18. Ciucanu, I. and F. Kerek, *A Simple and Rapid Method for the Permethylation of Carbohydrates.* Carbohydrate Research, 1984. **131**(2): p. 209-217.
19. Hayes, M.L. and F.J. Castellino, *Carbohydrate of the human plasminogen variants. III. Structure of the O-glycosidically linked oligosaccharide unit.* J Biol Chem, 1979. **254**(18): p. 8777-80.
20. Hortin, G.L., *Isolation of glycopeptides containing O-linked oligosaccharides by lectin affinity chromatography on jacalin-agarose.* Anal Biochem, 1990. **191**(2): p. 262-7.
21. Pirie-Shepherd, S.R., R.D. Stevens, N.L. Andon, J.J. Enghild, and S.V. Pizzo, *Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of human plasminogen 2.* J Biol Chem, 1997. **272**(11): p. 7408-11.
22. Ychou, M., J. Duffour, A. Kramar, S. Gourgou, and J. Grenier, *Clinical significance and prognostic value of CA72-4 compared with CEA and CA19-9 in patients with gastric cancer.* Dis Markers, 2000. **16**(3-4): p. 105-10.
23. Reiter, W., P. Stieber, C. Reuter, D. Nagel, C. Cramer, H. Pahl, and A. Fateh-Moghadam, *Prognostic value of preoperative serum levels of CEA, CA 19-9 and CA 72-4 in gastric carcinoma.* Anticancer Res, 1997. **17**(4B): p. 2903-6.
24. Werther, J.L., S. Rivera-MacMurray, H. Bruckner, M. Tatematsu, and S.H. Itzkowitz, *Mucin-associated sialosyl-Tn antigen expression in gastric cancer correlates with an adverse outcome.* Br J Cancer, 1994. **69**(3): p. 613-6.
25. Pinho, S., N.T. Marcos, B. Ferreira, A.S. Carvalho, M.J. Oliveira, F. Santos-Silva, A. Harduin-Lepers, and C.A. Reis, *Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells.* Cancer Lett, 2007. **249**(2): p. 157-70.

Chapter V. Discussion and Conclusion

The detection of sialylated oligosaccharides as well as sialylated glycoproteins have been widely studied. Several approaches have been presented, such as immunohistochemistry with antibodies [1] or lectins, but still structural validation is needed. In this context mass spectrometry is a high sensitive tool for the analysis of carbohydrates [2-4]. Since acidic glycans undergo metastable fragmentation in MALDI conditions [2-5], permethylation or similar derivatizing methods are often necessary to stabilize the sialic acids and unsure detection using this type of ionization [2, 6]. The third chapter of this work was dedicated to the development a method for the analysis of a particular class of sialylated O-glycans, the TF-related antigens, frequently (over)expressed in cancer and other pathologies. The method relies on the release of O-glycans from proteins immobilized in SDS-PAGE gels by reductive β-elimination followed by permethylation and nanoLC-MALDI-TOF/TOF analysis. The MS data generated in this way was then analyzed by a bioinformatics survey for the ions of interest and combined in a sparse plot matrix such as the one presented in Figure 2 - chapter III. The information provided by these plots information was found crucial to distinguish chromatographic envelopes related with the targeted ions. Moreover it provided means to select chromatographic runs for MS/MS analysis.

The proof of concept was achieved using fetuin, a widely studied glycoprotein, known to express sialyl-Tn, sialyl-3-T, sialyl-6-T and disialyl-T [7, 8]. All the above mentioned structures have been detected by our method having as starting material only low *pico*mole of fetuin. To our knowledge, this is the first time that such low amounts of glycoprotein could be successfully screened for sialylated TF-related antigens. Noteworthy, this method allowed to differentiate the sialyl-6-T from sialyl-3-T based on their characteristic product ion spectra (Figure 4b and 5b).

The developed analytical approach was then successfully applied to the identification of tumor associated carbohydrate antigen sialyl-Tn in a glycoprotein isolated from the urine of rats with chemically-induced high-grade bladder tumors. Moreover, it permitted the identification of the same epitope in plasminogen isolated from the serum of patients with gastritis and intestinal metaplasia, a gastric cancer precursor lesions. Even though the clinical value of these observations warrants more investigations the structural bases have been accurately and undoubtedly set.

References

1. Feizi, T., *Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens.* Nature, 1985. **314**(6006): p. 53-7.
2. Zaia, J., *Mass spectrometry of oligosaccharides.* Mass Spectrom Rev, 2004. **23**(3): p. 161-227.
3. Zaia, J., *Mass spectrometry and the emerging field of glycomics.* Chem Biol, 2008. **15**(9): p. 881-92.
4. Zaia, J., *Mass spectrometry and glycomics.* OMICS, 2010. **14**(4): p. 401-18.
5. Demelbauer, U.M., M. Zehl, A. Plematl, G. Allmaier, and A. Rizzi, *Determination of glycopeptide structures by multistage mass spectrometry with low-energy collision-induced dissociation: comparison of electrospray ionization quadrupole ion trap and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron time-of-flight approaches.* Rapid Commun Mass Spectrom, 2004. **18**(14): p. 1575-82.
6. Ciucanu, I., *Per-O-methylation reaction for structural analysis of carbohydrates by mass spectrometry.* Anal Chim Acta, 2006. **576**(2): p. 147-55.
7. Royle, L., T.S. Mattu, E. Hart, J.I. Langridge, A.H. Merry, N. Murphy, D.J. Harvey, R.A. Dwek, and P.M. Rudd, *An analytical and structural database provides a strategy for sequencing O-glycans from microgram quantities of glycoproteins.* Anal Biochem, 2002. **304**(1): p. 70-90.
8. Merry, A.H., D.C. Neville, L. Royle, B. Matthews, D.J. Harvey, R.A. Dwek, and P.M. Rudd, *Recovery of intact 2-aminobenzamide-labeled O-glycans released from glycoproteins by hydrazinolysis.* Anal Biochem, 2002. **304**(1): p. 91-9.