



Provided by the author(s) and NUI Galway in accordance with publisher policies. Please cite the published version when available.

Title	N- and O- linked Glycosylation, Developing Mass Spectrometric Strategies for the Characterisation of Glyco-epitopes
Author(s)	Kenny, Diarmuid T
Publication Date	2012-12-10
Item record	http://hdl.handle.net/10379/3281

Downloaded 2020-10-17T04:17:40Z

Some rights reserved. For more information, please see the item record link above.



N- and *O*- linked Glycosylation

Developing Mass Spectrometric Strategies
for the Characterisation of Glyco-epitopes

Diarmuid T. Kenny

Department of Chemistry
National University of Ireland, Galway

July 2012

Supervisors

Dr. Niclas G. Karlsson

Department of Medical Biochemistry
Institute of Biomedicine
University of Gothenburg

Prof. Robert J. Woods

Department of Chemistry
School of Science
National University of Ireland, Galway

Table Of Contents

1	Introduction	1
1.1	<i>N</i> - & <i>O</i> - linked oligosaccharides and sulfated glycosaminoglycans.....	2
	<i>N</i> -linked oligosaccharides	3
	<i>O</i> -linked oligosaccharides	5
	<i>Helicobacter pylori</i> and its interaction with the gastric mucin.....	8
	Sulfated oligosaccharides.....	10
1.2	Analytical approaches for the characterization of oligosaccharides	13
	Liquid chromatography	13
	Mass spectrometry	14
	Fragmentation of oligosaccharides	16
	Nomenclature for describing MS ⁿ fragmentation	18
	LC-MS analysis of oligosaccharides.....	18
1.3	MS analysis of oligosaccharides	24
	MS analysis of <i>N</i> - and <i>O</i> - linked oligosaccharides	24
	MS analyses of sulfated oligosaccharides.....	25
2	Aims of Thesis	26
3	Results and Discussion.....	27
3.1	Migration of sulfates	28
3.2	LC-MS of <i>O</i> - and <i>N</i> - linked oligosaccharides.....	30
	<i>N</i> -linked oligosaccharide characterization: A semi quantitative approach.....	30
	Characterization of MUC5AC derived <i>O</i> -linked oligosaccharides	32
3.3	Biological application of MS based glycomic analysis	35
	Identification of IacdiNAc: Possible implications in <i>Helicobacter pylori</i> infection?.....	35
	The importance of detecting aberrant glycosylation.	40
4	Conclusion	44
5	Future Perspectives	46
6	Supplementary Materials and Methods.....	47
7	Acknowledgements	51
8	References.....	53
9	Personal Involvement in Each Paper.....	62

Declaration of Originality

I, Diarmuid T. Kenny, hereby declare that this thesis submission is my own work. I have not obtained a degree at the National University of Ireland, Galway, or elsewhere, on the basis of this work. Any contribution made to the research by others with whom I have worked with is explicitly acknowledged in section 9 'Personal Involvement in Each Paper' of this thesis.

Abstract

The overall goal of this thesis was to adopt a liquid chromatography-mass spectrometry (LC-MS) platform for glycomic analysis of glycoproteins. In particular we were interested in characterizing complex glycomic samples such as membrane associated *N*-linked oligosaccharides and the mucin *O*-linked oligosaccharides with the aim of identifying particular glycosylation traits or glyco-epitopes that could have biological significance. For the manual interpretation of MS_n spectra of oligosaccharides, the acquisition of information-rich MS_n spectra is essential. The rearrangement of fragment ions during collision induced dissociation (CID) can complicate the interpretation of the MS_n and result in the mis-assignment of an oligosaccharide structure. We have shown that the migration of sulfate groups can occur when sulfated oligosaccharides are fragmented by CID in negative ion mode. The migration is promoted by the presence of a mobile proton and the steric availability of the sulfate groups. Fragmentation by high energy C-trap dissociation (HCD) limits the migration of sulfated residues to where it cannot be detected.

Our analysis of membrane associated *N*-linked oligosaccharides showed that the use of semi-quantitative data can be useful tool at highlighting difference in glycosylation and the use of statistical tools such as monosaccharide composition analysis (MSAC) was useful at providing a global overview of the glycosylation or sorting the oligosaccharide according to particular glycosylation traits such as common core extensions or terminal epitopes provided information on the biological aspects of the oligosaccharides. Our characterisation of *O*-linked oligosaccharides derived from MUC5AC of healthy and tumour-associated tissue revealed the presence of the novel lacdiNAc (GalNAcβ1-4GlcNAc) epitope on gastric mucin. We identified the lacdiNAc epitope by comparing the MS₃ of the lacdiNAc containing structure to known standards with terminal *N*-acetylgalactosamine (GalNAc) and *N*-acetylglucosamine (GlcNAc) residues and by digestion of the lacdiNAc containing oligosaccharides with *exo-N*-hexosaminidase. We believe this epitope plays a role in inhibiting the binding and proliferation of *Helicobacter pylori* to MUC5AC by limiting the synthesis of oligosaccharides displaying Leb and sLex epitopes which are required for the adhesion of *H. pylori* via its adhesins BabA and SabA respectively.

List of Papers

This thesis is based on the following papers which are referred to in the text by their Roman numerals

- I Kenny, D.T., C.A. Hayes, C. Jin, and N.G. Karlsson, Perspective and Review of Mass Spectrometric Based Sulfoglycomics of *N*-Linked and *O*-Linked Oligosaccharides. *Current Proteomics*, 2011. **8**(4): p. 278-296
- II Kenny, D.T., S.M.A. Issa, and N.G. Karlsson, Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation. *Rapid Communications in Mass Spectrometry*, 2011. **25**(18): p. 2611-2618.
- III Kenny, D.T., L. Ali, S.M.A Issa and N. G. Karlsson, Glycomic Analysis of Membrane-Associated Proteins, in *Sample Preparation in Biological Mass Spectrometry*, A.R. Ivanov and A.V. Lazarev, Editors. 2011, Springer. p. 498-513.
- IV Kenny, D.T., K.A. Thomsson, and N. G. Karlsson, Semi-Quantitative Data Analysis for Membrane Associated *N*-linked Oligosaccharides Highlights Differences in the Glycosylation of Two Enriched Membranes Samples. *Manuscript*
- V Kenny, D.T., E.C. Skoog, S.K. Linén, W.B. Struwe, P.M. Rudd and N.G. Karlsson Presence of Terminal *N*-acetylgalactosamine β 1-4*N*-acetylglucosamine Residues on *O*-linked Oligosaccharides from Gastric MUC5AC: Involvement in *Helicobacter pylori* Colonization? *Glycobiology*, 2012. **22**(8): p. 1077–1085

All papers were reproduced with the permission of the copyright holder.

Abbreviations

Asp	Asparagine
BabA	Blood group antigen binding Adhesin
CID	Collision induced dissociation
CRT	Calreticulin
CS	Chondroitin sulfate
DHB	2,5-dihydroxybenzoic acid
DS	Dermatan sulfate
ESI	Electrospray ionization
Fuc	Fucose
Gal	Galactose
GalNAc	<i>N</i> -acetylgalactosamine
GalNAc-ol	<i>N</i> -acetylgalactosaminitol
GDP	Guanosine-diphosphate
GlcA	Glucuronic acid
GlcNAc	<i>N</i> -acetylhexosaminitol
GPI	Glycylphosphatidylinositol
HCD	Higher C trap dissociation
Hex	Hexose
HexNAc	<i>N</i> -acetylhexosamine
HGM	Human gastric mucin
HPLC	High performance liquid chromatography
<i>H. Pylori</i>	<i>Helicobacter pylori</i>
HSA	High sialic acid membrane
IdoA	Iduronic acid
ILM	Ionic liquid matrix
KS	Keratan sulfate
LacdiNAc	Gal β 1-4GlcNAc
LacNAc	<i>N</i> -acetylactosamine
LC-MS	Liquid chromatography-mass spectrometry
LDH	Lactate dehydrogenase
Le ^a	Lewis A
Le ^b	Lewis B
Le ^Y	Lewis Y
Le ^x	Lewis X
LSA	Low sialic acid membrane
MALDI	Matrix assisted laser desorption ionization
Man	Mannose
MS	Mass spectrometry
MS ⁿ	Tandem mass spectrometry
<i>m/z</i>	Mass to charge
NaBH ₄	Sodium borohydride
NaOH	Sodium hydroxide
Neu5Ac	<i>N</i> -acetylneuraminic acid (Sialic acid)
Neu5GC	<i>N</i> -glycolylneuraminic acid (Sialic acid)
NMR	Nuclear magnetic resonance
PGC	Porous graphitized carbon
PGM	Porcine gastric mucin

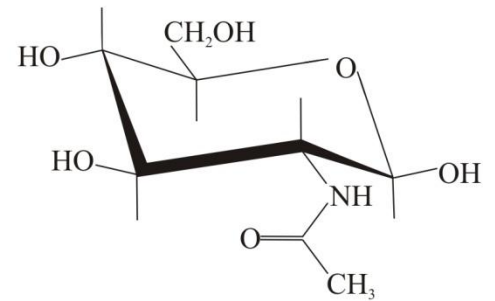
Abbreviations

PNGase F	Peptide: N-glycosidase F
PVDF	Polyvinylidene fluoride
QIT	Quadrapole ion trap
SabA	Sialic acid binding Adhesin
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Thr	Threonine
TOF	Time of flight
UDP	Uridine-diphosphate
VNTR	Variable number of tandem repeats
Xyl	Xylose

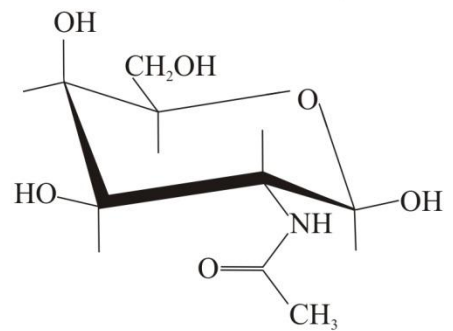
Monosaccharides Were Represented According to the Consortium for Functional Glycomics Notation



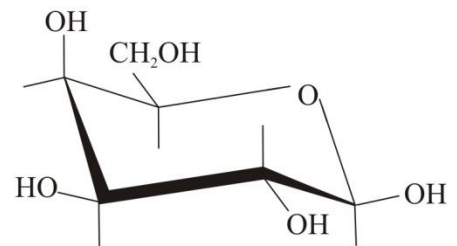
β -D-N-Acetylglucosamine



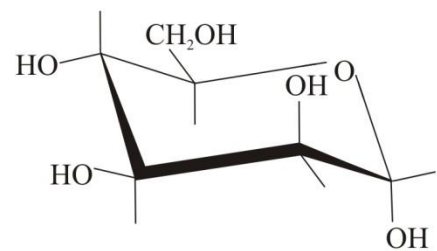
β -D-N-Acetylgalactosamine



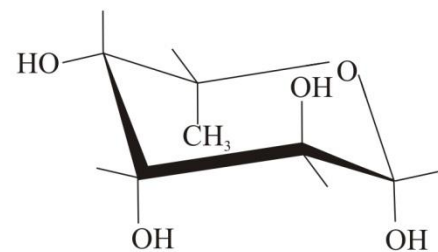
β -D-Galactose



α -D-Mannose

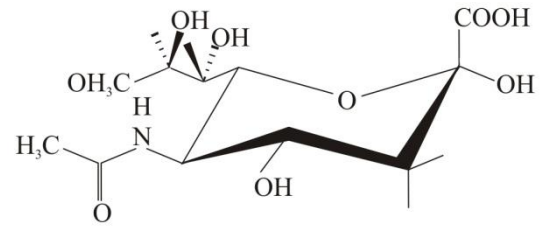


α -L-Fucose

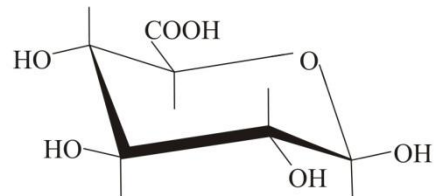




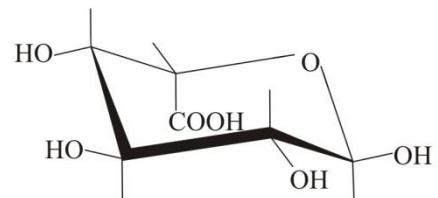
α -D-N-acetylneuraminic Acid



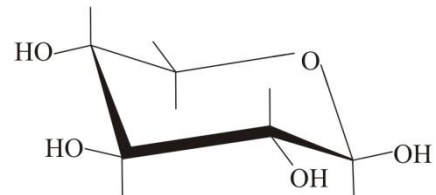
β -D-Glucuronic Acid



β -D-Iduronic Acid



β -D-Xylose



Note: A proton is present for the linkages where no functional group is displayed



Sulfate



Reducing Aldose



Y_x or C_x fragments



Reduced Alditol



Z_x or B_x fragments



Cross-Ring Fragmentation

1 Introduction

Proteins undergo various modifications during biosynthesis. Glycosylation, one such modification, involves the covalent bonding of carbohydrates (oligosaccharides) to either the nascent polypeptide (co-translation) or the fully formed protein (post-translation)¹. In mammals, nearly half of all proteins as well as many lipids undergo glycosylation². The carbohydrate portion of a glycoprotein can be as large as the protein or even surpass the protein content. However, the importance of carbohydrates in biological function of a glycoprotein has often been overlooked. There is an enormous diversity with respect to the glycosylation of a protein, both at the macro- and micro- level. Glycosylation does not always occur on every oligosaccharide receptor site and different oligosaccharides can occupy the same oligosaccharide receptor site for a particular protein³. Considering the ubiquitous nature of glycosylation, it is not surprising that it plays a considerable role within many cellular functions from development and growth to the proper functioning and survival of the cell⁴. The biological role of oligosaccharides can broadly be subdivided into two groups: Some oligosaccharides play an important role in the structural properties of a glycoprotein; for example the proteoglycan aggrecan plays integral role in providing structural support in cartilage tissue⁵ and the mucin oligosaccharides found on the mucosal surfaces similarly play a structural role in providing lubrication and protection to the epithelial surface⁶. Other Oligosaccharides provide specific recognition properties that are mediated by particular receptors such as the blood groups involvement in transporting molecules through membranes⁷ or the Lewis epitopes role in antigen binding⁸.

Glycomics, the comprehensive study of the carbohydrate content of an organism (glycome), includes the characterization of the complex oligosaccharides present on glycoproteins. The analysis of complex carbohydrates can be challenging as unlike protein synthesis, glycosylation does not have a highly regulated biosynthetic pathway with a predefined code. Whereas proteins follow a linear biosynthetic pathway, whereby DNA undergoes a tightly regulated process of transcription and translation to form the protein, the biosynthetic pathway for glycosylation is dynamically regulated by reactions in the endoplasmic reticulum (ER) and Golgi apparatus². These reactions can be affected by extra- and intra- cellular stimuli and it can thus be impossible to predict the final glycosylation profile of a particular cell due to the lack of a 'code' akin to DNA for oligosaccharides. This results in a mix of different glyco-epitopes that are elaborated around common core structures. Alteration in the glycosylation profile can occur which may have a variety of biological consequences; amongst these are several types of congenital muscular dystrophy caused by misglycosylation of dystroglycan^{9,10} as well as congenital disorders of glycosylation which can be attributed to a number of different misglycosylation events^{11,12}. Changes in glycosylation can be used as biomarkers to detect different types of cancers; for example over-sialylation of the stomach lining is associated with gastric cancer¹³.

Improper glycosylation can have a negative impact on glycosylated therapeutics where incomplete or misglycosylation can impact its efficacy or decrease its *in-vivo* half-life^{14,15}. With the high costs associated with biopharmaceutical products, the biotechnology industry is naturally very interested in understanding and defining the glycosylation profile of a glycoprotein to ensure there is oligosaccharide homogeneity between fermentation batches^{14,16}.

Although this thesis will focus on *N*-linked and *O*-linked glycosylation with references to the analysis of the sulfated glycosaminoglycan's (GAG's)^{17,18} (see figure 1), glycomics also includes the study of the carbohydrate content of glycolipids^{19,20}, GPI anchors^{21,22} and C-mannosylation²³ amongst others.

1.1 *N*- & *O*- linked oligosaccharides and sulfated glycosaminoglycans

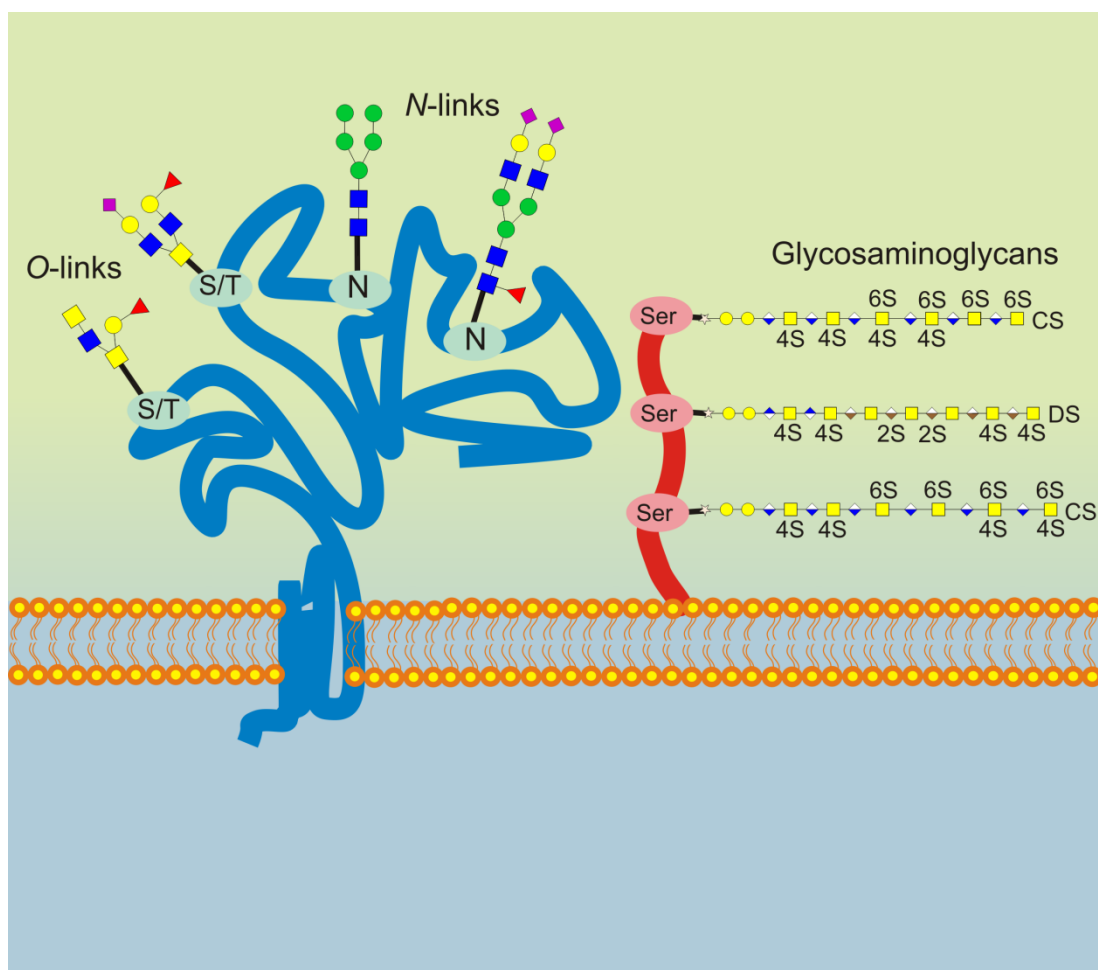


Figure 1: *N*- and *O*- linked oligosaccharide attached to a membrane associated glycoprotein and a proteoglycan with chondroitin sulfate and dermatan sulfate. The *N*-linked oligosaccharides are attached via a nitrogen residue within the consensus sequon Asp-Xxx-Ser/Thr (where Xxx is not Pro). The *O*-linked oligosaccharides are attached via a hydroxyl oxygen of a Ser or Thr residue. Chondroitin sulfate and dermatan sulfate are attached via a hydroxyl oxygen on a SER residue.

Monosaccharides are the simplest form of a sugar and the basic building block of all oligosaccharides. They are either aldehydes (-CHO functional group) or ketones (-C=O functional group) with multiple hydroxyl groups and have a chemical formula of $C_x(H_{2O})_y$. In general, mammalian glycoprotein oligosaccharides can be described as a combination of hexoses (mannose (Man), galactose (Gal) and glucose (Glu)), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), fucose (Fuc), uronic acids (glucuronic (GlcA) and iduronic acid (IdoA)), xylose (Xyl) and sialic acids (most commonly *N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc))².

Biosynthesis of oligosaccharides is initiated in the endoplasmic reticulum and proceeds in the Golgi apparatus by the actions of ER- and Golgi apparatus resident glycosyltransferases (as well as the action of some glycosidases). Generally, glycosyltransferases will sequentially add a specific monosaccharide residue via a specific linkage position to a defined acceptor site thus creating linear or branched structures. This has led to the general hypothesis of ‘one-enzyme, one linkage’ hypothesis for the biosynthetic process, whereby families of glycosyltransferases exist with each member of a family catalyzing a specific reaction between a donor and acceptor². The monosaccharide that is being transferred to an acceptor site is present in the ER-Golgi pathway as an activated nucleotide sugar. The majority of the nucleotide donor sugars are present as uridine-diphosphates (UDP) and are thus called UDP-Gal, UDP-GlcNAc etc. except for Fuc and Man residues which are present as guanosine-diphosphate (GDP) sugars and Neu5Ac which is present as cytidine-monophosphate (CMP) sugar.

***N*-linked oligosaccharides**

N-linked glycosylation involves the covalent bonding of oligosaccharides onto the nitrogen of an asparagine residue within the consensus sequon Asp-Xxx-Ser/Thr (where Xxx is not Pro)²⁴. Although this consensus sequon is necessary for a particular site on the peptide to be *N*-linked glycosylated, the presence of this sequon does not necessarily imply that it will be glycosylated²⁵. Conformational or other constraints during protein folding may restrict the potential *N*-linked glycosylation site from accepting an oligosaccharide^{26,27}. *N*-linked glycosylation is a co-translation event, whereby the oligosaccharides are attached to a nascent polypeptide. All *N*-linked oligosaccharides share a common GlcNAc₂Man₃ core sequence. However since the biologically relevant part of an *N*-linked oligosaccharide is usually towards the terminal portion of the oligosaccharide, it is common to describe them by their terminal characteristics. Generally *N*-linked oligosaccharides can be categorized as one of three structural types: high mannose, complex or hybrid (see figure 2).

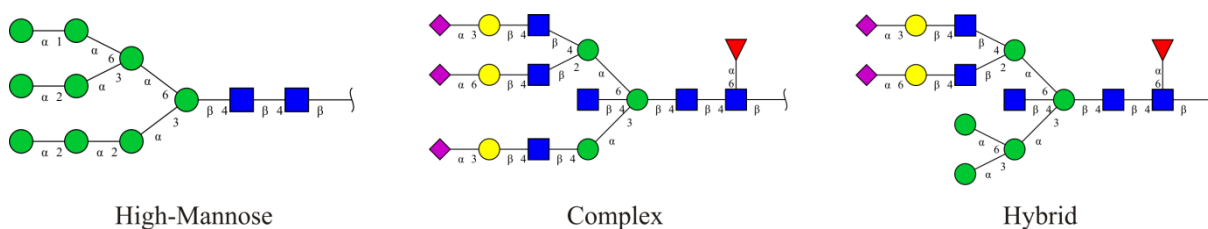


Figure 2: High-Mannose, Complex and Hybrid *N*-linked oligosaccharides

Most secreted as well as membrane bound proteins are *N*-linked glycosylated²⁸ and the oligosaccharides play a role in many cellular functions. Some *N*-linked oligosaccharides contain epitopes that bind to lectins²⁹ or to other carbohydrates³⁰, other *N*-linked oligosaccharides are involved in cell to cell adhesion³¹⁻³³. Furthermore *N*-linked oligosaccharides can act as self-recognition epitopes for the cells immune system³⁴. *N*-linked oligosaccharides are also important for proper folding of proteins during biosynthesis^{27,35}.

Changes in the *N*-linked glycosylation can have profound effects on the function of the glycoprotein. Terminal sialylation, in particular, can play a primary role in the function of some *N*-linked oligosaccharides^{36,37}. Determining the level of sialylation is important both clinically, where changes in sialylation have been associated with various forms of cancer³⁸⁻⁴⁰, and commercially where the decrease in the level of sialylation can negatively impact the efficacy of a biopharmaceutical^{14,34}. The under-sialylation of glycosylated biopharmaceuticals can occur through the introduction of various waste products into the culture media during the fermentation process. For example ammonium chloride, a byproduct of glutamine metabolism, can alter the inter-compartmental pH of the Golgi apparatus and cause redistribution of the glycosyltransferases⁴¹. Since the transfer of sialic acid occurs in the distal compartments of the Golgi apparatus shortly before the glycoprotein is secreted, the redistribution of the glycosyltransferases can inhibit the activity of the resident sialyltransferases⁴². This will result in the secretion of under- or asialylo- glycoproteins. In our analysis of a fusion glycoprotein secreted from a Chinese hamster ovary (CHO) cell line, we observed that improper glycosylation occurs during the end of a batch fermentation cycle where increased concentrations of ammonium chloride would be expected (See Discussion, section 3.3).

Biosynthesis of *N*-linked oligosaccharides

Biosynthesis of *N*-linked oligosaccharides begins with the synthesis of Glc₃Man₉GlcNAc₂ intermediary which is transferred *en-bloc* to a target asparagine residue of the nascent polypeptide. Once transferred to the protein, it undergoes a process of trimming by various glycosidases to produce a high-mannose Man₅GlcNAc₂-Asn. This is eventually processed to form core the Man₃GlcNAc₂ structure that is common to all *N*-linked oligosaccharides. An important function of *N*-linked glycosylation is to ensure that proper folding of the protein occurs⁴³. This quality control process begins in the ER with interaction of a monoglucosylated intermediary, GlcMan₉GlcNAc₂, with two resident chaperone proteins ; the membrane-bound calnexin (CNX) and the soluble calreticulin (CRT)⁴⁴⁻⁴⁶ as well as the disulphide isomerase ERp57 which forms transient disulphide bonds with both CNX and CRT. This facilitates the folding of the glycoprotein into its native form⁴⁷. Once proper folding has been achieved, the last remaining Glc residue from the complex is removed, which releases the properly folded protein and this allows it to enter the Golgi apparatus for further processing^{47,48}. If a protein is continually mis-folded, it gets relocated to the cytosol where it is eventually degraded⁴⁹.

***O*-linked oligosaccharides**

In *O*-linked glycosylation the oligosaccharides are covalently bound to the hydroxyl oxygen of an amino acid on a protein. In contrast to the co-translation process of *N*-linked glycosylation, where an oligosaccharide is transferred *en-bloc* to a nascent polypeptide, *O*-linked glycosylation is a post translation modification that begins with the addition of a single monosaccharide to a fully formed protein. There are different types of *O*-linked oligosaccharides⁵⁰⁻⁵³ however the mucin type GalNAc *O*-linked glycosylation⁵⁴ is perhaps the best understood. In mucin-type *O*-linked glycosylation, the oligosaccharides are attached to either a Ser or Thr residue^{2,55}.

Mucins are high molecular weight gel-forming glycoproteins and are a major component of the bodies mucus secretions. Mucins are encoded by the MUC family of genes, of which over 20 different MUC genes have been identified⁵⁶. Although the biosynthetic pathway of mucin type *O*-linked oligosaccharides (henceforth referred to as *O*-linked oligosaccharides) has been extensively studied and many aspects of *O*-linked glycosylation are well understood⁵⁷, no known consensus sequence has been discovered⁵⁴.

Mucins are heavily *O*-linked glycosylated, a distinguishing feature which is derived from the presence of a repeated peptide sequence called 'variable number of tandem repeats' (VNTR). VNTR's are rich in both Ser and Thr residues which gives the mucin a large number of *O*-linked glycosylation acceptor site⁵⁸. The *O*-linked oligosaccharides play an important role in the function of mucin glycoprotein from their intrinsic role in the ABO⁵⁹ and Lewis blood groups⁶⁰ to its protective role as a major constituent of the glycocalyx that covers the secretory epithelial layers such as the respiratory tract, gastro-intestinal tract and ocular surface^{61, 62}.

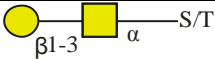
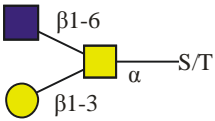
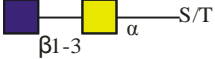
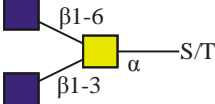
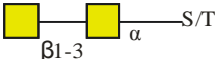
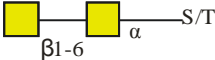
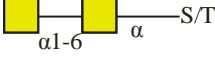

Biosynthesis of *O*-linked oligosaccharides

In the human mucin, oligosaccharides are predominantly made up of five monosaccharide residues; Gal, GalNAc, GlcNAc, Fuc and the sialic acid Neu5Ac. The mucin can be further modified with the addition of sulfate groups which is usually found on either the Gal or GlcNAc residues⁵⁷. *O*-linked glycosylation is initiated by the addition of a single α linked GalNAc to the hydroxyl group of an acceptor Ser or Thr residue by the glycosyltransferase GalNAc-T⁶³. Twenty different GalNAc-T genes have been identified in human. GalNAc-T's are expressed in a tissue-specific manner and show a preference for different peptide sequences^{63,64}.

Since *O*-linked glycosylation is a post translation event, it is assumed that only the Ser or Thr residues that are exposed on the surface of a folded protein are accessible for glycosylation. Furthermore, it has been observed that mucin *O*-linked glycosylation receptor sites are usually present on the β -turns of the polypeptide backbone and on regions with extended conformation. It is also rare to find larger amino acids such as tyrosine or tryptophan in the vicinity of an *O*-linked glycosylation receptor site as this steric hindrance would limit the accessibility of the GalNAc-T enzymes⁵⁷.

Most mucin oligosaccharides can be divided into three constituent parts: 1) A core unit linked to the protein 2) an elongated backbone chain and 3) terminating epitopes. Eight different *O*-linked core structures have been identified⁵⁷ (See table 1). Our analysis of the gastric mucin revealed that the majority of the oligosaccharides identified on MUC5AC were core 1, 2 and 3 with smaller number of core 4 structures (Paper V). The simplest core structures are the disaccharide core 1 and core 3. Core 1 is formed by the addition of a Gal β 1-3 to the α -linked Gal and this is a common event in mucin and non-mucin *O*-linked glycosylation, however the formation of core 3 by addition of a GlcNAc β 1-3 seems to be specific to mucins. The trisaccharide core 2 and core 4 are formed by the addition of a β 1-6GlcNAc to the core 1 and core 3 structures respectively by various different β 1-6GlcNAc transferases. Extension of the core structures occurs by the addition of either a Type 1 (Gal β 1-3GlcNAc β 1) or Type 2 (Gal β 1-4GlcNAc β 1) *N*-Acetylglucosamine unit.

Table 1. Table of Mucin *O*-linked oligosaccharide Core structures

Core	Structure	
Core 1	Gal β 1-3GalNAc α 1-Ser/Thr	
Core 2	Gal β 1-3(GlcNAc β 1-6)GalNAc α 1-Ser/Thr	
Core 3	GlcNAc β 1-3GalNAc α 1-Ser/Thr	
Core 4	GlcNAc β 1-3(GlcNAc β 1-6)GalNAc α 1-Ser/Thr	
Core 5	GalNAc β 1-3GalNAc α 1-Ser/Thr	
Core 6	GalNAc β 1-6GalNAc α 1-Ser/Thr	
Core 7	GalNAc α 1-6GalNAc α 1-Ser/Thr	
Core 8	Gal α 1-3GalNAc α 1-Ser/Thr	

Mucin oligosaccharides have been found to display the ABH and the Lewis blood group epitopes. Type 1 (Gal β 1-3GlcNAc β 1-) extension of the core structure can terminate with Le^a, Le^b, sLe^a as well as Type 1 ABH blood group antigens and Type 2 (Gal β 1-4GlcNAc β 1) extended core structures can display Le^x, Le^y, sLe^x as well as the Type 2 ABH blood group antigens (Table 2). The activity of the secretor gene (α 2 fucosyltransferase) is a prerequisite for the expression of the blood group antigens, thus they will not be present on *O*-linked oligosaccharides from non-secretor individuals.

The secretor status of an individual will therefore influence the mucin glycosylation profile. The display of the Lewis blood groups on the gastric mucins MUC5AC and MUC6 of secretor individuals in particular can influence their susceptibility to infection by certain microbes that can utilise the Lewis epitopes as functional binding sites.

Table 2. *O*-linked oligosaccharide elongation and termination structures

Nomenclature	Structure
<i>N</i>-Acetyllactosamine	
Type 1	-Gal β 1-3GlcNAc β 1-
Type 2	-Gal β 1-4GlcNAc β 1-
Terminal structures	
Blood group H	Fuca1-2Gal β 1-
Blood group A	Fuca1-2(GalNAca1-3)Gal β 1-
Blood group B	Fuca1-2(Gal α 1-3)Gal β 1-
Terminal structures	
Type 1	
Lewis a (Le^a)	Gal β 1-3(fuca1-4)GlcNAc β 1-
Lewis b (Le^b)	Fuca1-2Gal β 1-3(Fuca-4)GlcNAc β -
Sialyl-Le^a	NeuAc(α 2-3)Gal β 1-3(Fuca1-4)GlcNAc β 1-
Type 2	
Lewis x (Le^x)	Gal β 1-4(Fuca1-3)GlcNAc β 1-
Lewis y (Le^y)	Fuca1-2Gal β 1-4(Fuca1-3)GlcNAc β 1- (include H)
Sialyl-Le^x	NeuAca2-3Gal β 1-4(Fuca1-3)GlcNAc β 1-
Sulfation	
3 Sulfation	HSO ₃ -3Gal β 1-
6 Sulfation	HSO ₃ -6GlcNAc β 1-
Examples of combined epitopes	
H-Type 2	Fuca1-2Gal β 1-4GlcNAc β -
Sialylated Type 2	NeuAca2-3Gal β 1-4GlcNAc β 1-

The glycosylation profile of mucins changes dramatically with the onset of cancer and specific *O*-linked structures have been associated with cancer. The sole addition of an α -linked GalNAc to a protein creates the Tn antigen⁶⁵. This can be elongated with the addition of a α 2-6Neu5Ac which results in the sialyl-Tn (sTn) antigen^{65,66} or the addition of a β 1-3Gal to produce the T (also known as the TF or Thomsen-Friedenreich) antigen⁶⁷. These antigens are not usually found in healthy tissue however they have been associated with cancer where incomplete glycosylation is a common event⁶⁸. However, these epitopes can be displayed in non-pathological circumstances. In our analysis of *O*-linked oligosaccharides from MUC5AC from healthy and cancerous stomachs, we detected oligosaccharides displaying the Tn, sTn and T epitopes on tumor associated tissue but not on the healthy tissue (Paper V). The primary sialic acid in humans is Neu5Ac and is usually present on *O*-linked oligosaccharides as α 2-3 (as in sL^{a/x}) or α 2-6 (as in sTn) linked³⁴. The sialic acid Neu5Gc has been found in cancer tissue and has been identified as a possible biomarker for cancer⁶⁹.

***Helicobacter pylori* and its interaction with the gastric mucin**

Mucin glycoproteins are a major constituent of the mucous layer that lines the gastrointestinal tract. In the stomach, the mucus is constantly being recycled and this recycling process protects the underlying epithelial cells from microbiological attack as invading microbes are discarded during the renewal process. In the gastrointestinal tract the mucus layer protects the epithelial cells from the harsh pH environment maintained by the stomach while simultaneously allowing for nutrient and water uptake^{6,70}. However, the mucosal epitopes can also be utilised by some invading microbes that can utilize them as adhesion sites. *Helicobacter pylori*, a spiral-shaped gram-negative bacterium, is one such microbe that can colonize the gastric epithelium. *H. pylori* was first isolated by Marshall and Warren in 1984 and this discovery revealed *H. Pylori*'s association with gastritis, peptic ulcers and several forms of gastric cancer⁷¹. *H. pylori* is the most common bacterial infection in the world^{72,73} and can infect up to 90% of the population in developing countries and up to 50% in developed countries⁷⁴.

Attachment of *H. pylori* to the gastric surface is mediated by lectin-like outer membrane proteins known as adhesins present on *H. pylori* as well as the presence of specific blood group epitopes on the gastric mucin. Two adhesins in particular have been identified as being important in the adhesion process; The Blood group antigen binding Adhesin (BabA) which binds to Le^b (Fuc α 1-2Gal β 1-3(Fuc α -4)GlcNAc β -) and blood group H Type 1 (Fuc α 1-2Gal β 1-3) epitopes^{75,76} and Sialyl antigen binding Adhesin (SabA) which binds to sLe^x (NeuAc α 2-Gal β 1-4(Fuc α 1-3)GlcNAc β 1-) epitopes⁷⁷.

The primary secreted mucins lining the gastric epithelial layer are MUC5AC and MUC6. In the early onset stage of *H. pylori* infection, the mucins primarily express neutral epitopes, thus BabA-mediated binding of *H. pylori* to the mucus membranes via Le^b structures is the primary mode of adhesion^{78,79}.

As the infection progresses and inflammation and/or cancer develop, the glycosylation profile of the mucus changes and the acidic sialylated and sulfated structures become more prominent. At this stage binding to the sialylated sLe^x structures via SabA also occurs⁷⁷. Chronic infection can eventually causes several pathological conditions through excessive inflammation of the mucosa (Figure 3)

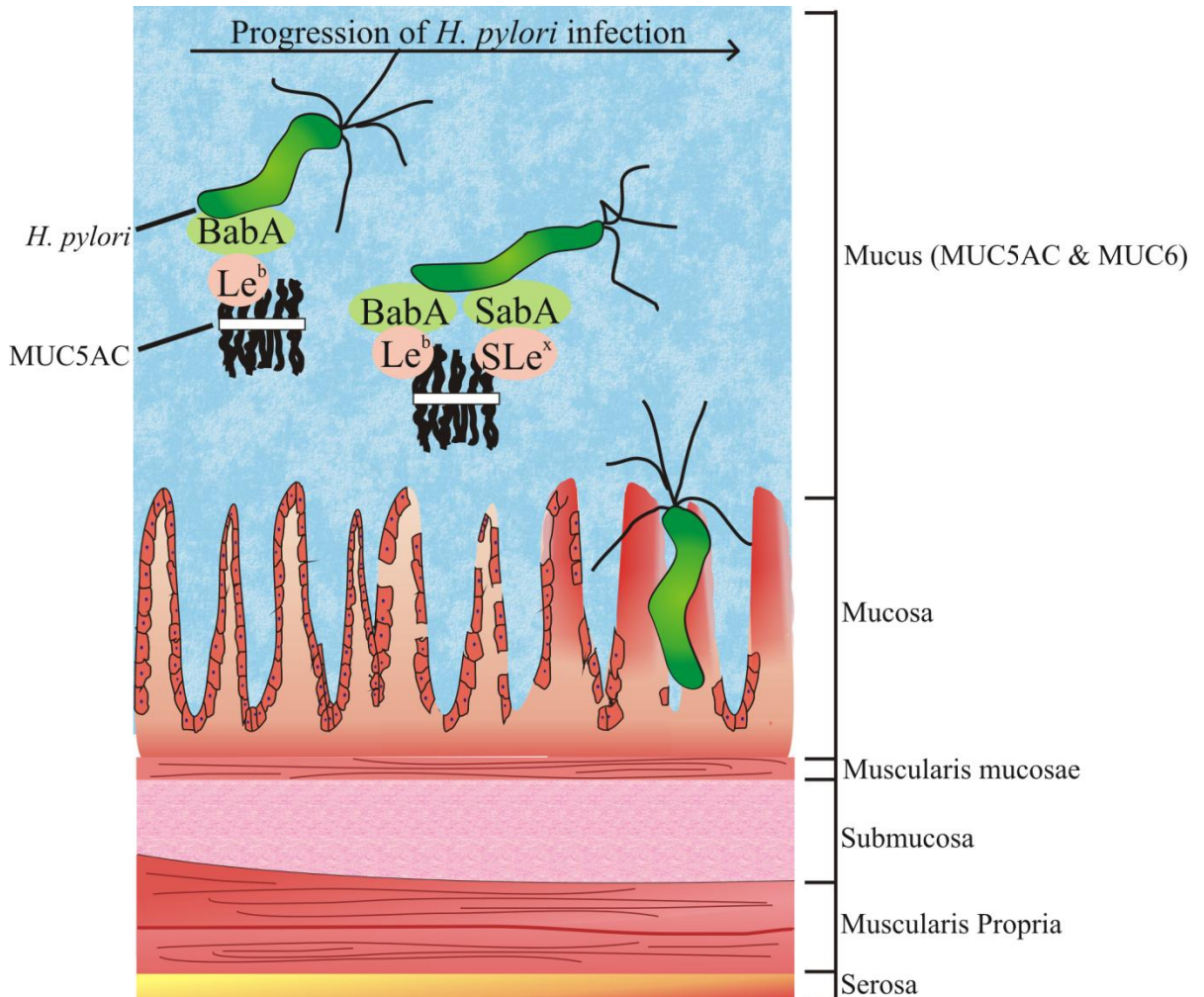


Figure 3: The progression of *H. pylori* infection in the stomach begins with the interaction of the lectin-like adhesin BabA to Le^b expressing oligosaccharides displayed on the gastric mucus. Inflammation of the gastric mucosa alters the glycosylation profile of the mucus and binding via SabA to sLe^x expressing oligosaccharides becomes more prominent. Chronic infection of *H. pylori* can eventually lead to gastritis, peptic ulcers and several forms of gastric cancer.

While the presence of Le^b and sLe^x containing structures have been associated with the promotion and binding of *H. pylori*, other epitopes such as GlcNAcα1-4Gal has been identified as a possible anti-bacterial epitope that suppresses the proliferation of *H. pylori* when present on the gastric mucus⁸⁰. The GlcNAcα1-4Gal- containing oligosaccharide were initially identified on MUC6 within the deeper portion of the gastric mucosa.

Recently, we have identified lacdiNAc (GalNAc β 1-4GlcNAc-) containing oligosaccharides on MUC5AC that may also be involved in the interaction between *H. pylori* and the mucous layer (Paper V). We identified a mucin isolated from a secretor individual with high levels of lacdiNAc containing oligosaccharides that inhibited the binding of *H. pylori*. Our initial investigation suggests that increased amount of the lacdiNAc epitope influences the expression of Le^b and complex sialylated structures. The complex relationship between *H. pylori* infection and the gastric mucosa clearly underpins the importance of developing analytical platforms that can characterize the oligosaccharide structure. Characterization of the oligosaccharide structures is integral to understanding its biological function.

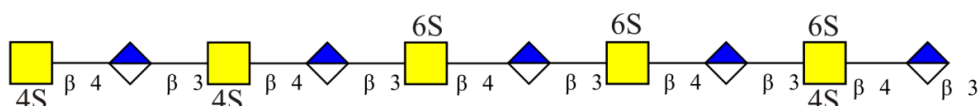
Sulfated oligosaccharides

A common modification of oligosaccharides is the addition of sulfate groups and their addition provides functionally specific negative charges to the oligosaccharide. The process by which a sulfate residue is added to the oligosaccharide is mediated by specific sulfotransferases, which are resident in the lumen of the Golgi apparatus, and transfer the sulfate from coenzyme-3'-phosphoadenosine-5'-phosphosulfate (PAPS) to the particular oligosaccharide that is to be sulfated⁸¹. Similar to the manner in which a specific glycosyltransferase is required to add a particular monosaccharide via a specific linkage during oligosaccharide biosynthesis, a specific sulfotransferase is required for each particular sulfate position and will recognise a specific monosaccharide structure, anomeric configuration or linkage position. Sulfotransferases can be specific for a particular group of glycoconjugates e.g. *N*-linked, *O*-linked, glycosaminoglycans (linear polysaccharides of repeating disaccharide units that contain either a GalNAc or GlcNAc residue attached to a uronic acid, either glucuronic acid (GlcA) or idouronic acid (IdoA)) or glycolipids etc. Sulfotransferase can also be less specific and can act where the precursor structure share a commonality regardless of conjugate type⁸¹. For example the same sulfotransferases may be involved in the sulfation of *N*- and *O*- linked oligosaccharides as well as the glycosaminoglycan keratan sulfate⁸² as they share the structural characteristic of lacNAc extension. The biosynthesis of sulfated oligosaccharides is a hierarchical event in that the backbone of the oligosaccharide is synthesized first, with the attachment of the sulfate to the oligosaccharide a secondary event. As a result of this hierarchical process, the localisation of a particular sulfotransferase in the Golgi apparatus can alter its activity in the biosynthetic pathway⁸¹. This is evident during the biosynthesis of a 3'-sulfo Le^x structure as the sulfate residue must be transferred to the oligosaccharide prior to fucosylation as the sulfotransferase is unable to act on a fucosylated precursor⁸³.

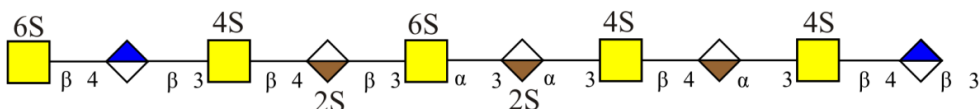
Proteoglycans

Sulfation of oligosaccharides is most commonly associated with the glycosaminoglycans (GAGs) of the proteoglycan family. Proteoglycans consist of a core protein with one or more covalently attached GAG chains. There are three different types of sulfated proteoglycans: 1) chondroitin sulfate (CS)/dermatan sulfate (DS), 2) heparan sulfate (HS)/heparin and 3) keratan sulfate (KS) (Figure 4)

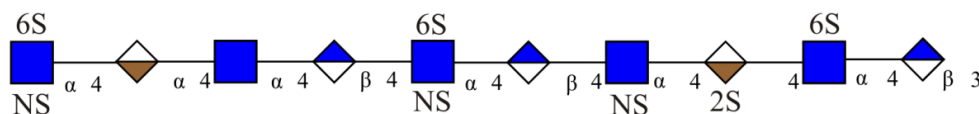
A Chondroitin Sulfate (CS)



B Dermatan Sulfate (DS)



C Heparan Sulfate (HS)/Heparin



D Keratan Sulfate (KS)

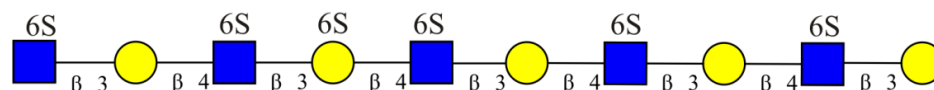


Figure 4: Typical glycosaminoglycans are long unbranched polysaccharides with repeating disaccharide units. A) Chondroitin sulfate B) Dermatan sulfate C) Heparan sulfate (with heparin) and D) Keratan sulfate

CS consists of monosulfated repeating disaccharide units of GalNAc β 1-4GlcA β 1-3. When the sulfate resides on the C-4 position of the GalNAc it is known as chondroitin-4-sulfate (CS-4) and when it is on the C-6 position it is known as chondroitin-6-sulfate (CS-6). Several variants of CS exist, these include dermatan sulfate (DS) which contains some IdoA residues and two oversulfated versions known as CS-D (GlcA2S-GalNAc6S) and CS-E (GlcA-GalNAc4S/6S).

Both CS-4 and CS-6 are major components of the proteoglycan aggrecan which is involved in the formation and maintenance of cartilage tissue⁸⁴. CS-4 has a wide variety of biological roles from lymphoid cell activation via the CD44 ligand⁸⁵ to the protection of high density lipoprotein against copper-dependent oxidation⁸⁶ while CS-6 is component of PTP ζ , which is a member of the protein-tyrosine phosphatase family of enzymes⁸⁷.

HS consists of variably sulfated GlcNAc α 1-4GlcA β 1-4 disaccharides. HS can also be oversulfated, in which case it is known as heparin⁸⁸. A variant of the GlcNAc α 1-4GlcA β 1-4 disaccharide includes a deacetylated version, GlcNS, which is usually present as IdoA2S-GlcNS6S. HS is a ubiquitous GAG produced in almost every cell whereas heparin is exclusively produced in mast cells. They are both involved in various biological processes including cell proliferation⁸⁹, differentiation⁹⁰ and adhesion⁹¹. HS is of particular medical importance since it is highly sulfated heparin version is commonly used as an anti-coagulant⁹².

KS shares structural affinity with both *N*- and *O*-linked glycosylation in so far as all three undergo poly-*N*-acetylactosamine (Gal β 1-4GlcNAc) extensions with sulfation on the C6 position of Gal and GlcNAc residues. The level of sulfation is mediated by the sulfation of Gal which always occurs after the GlcNAc residue has been sulfated. KS is mainly found in the cornea and cartilage tissue⁹³.

Sulfation of *O*- and *N*- linked oligosaccharides

Sulfation of *N*- and *O*- linked oligosaccharides is reported far more rarely in literature compared to GAG's. Nonetheless the transfer of a sulfate residue to *N*- and *O*- linked oligosaccharides is a ubiquitous event in their biosynthetic pathway. One of the earliest known roles of the sulfate on a carbohydrate was its involvement in the luteinizing hormone (LH) where sulfation of *N*-linked oligosaccharides are required for effective clearance of LH from the bloodstream⁹⁴. Activation of human natural killer cells (the sulfated HNK-epitope)⁹⁵ is also dependent on the sulfation of the *N*-linked oligosaccharides.

O-linked sulfation has been associated with the homing of lymphocytes to peripheral lymph nodes whereby special endothelium called high endothelial venules are known to contain sulfated oligosaccharides via the 6-sulfo sLe^x epitope⁹⁶. *O*-linked sulfation is a common modification of the mucin oligosaccharides and as a result has been a topical subject when discussing the mucin lining of the respiratory tract with relation to cystic fibrosis where it has been shown that the level of sulfation can be influenced by pulmonary bacterial infections and inflammation^{97,98}.

1.2 Analytical approaches for the characterization of oligosaccharides

Glycomics, the comprehensive study of the carbohydrate content (glycome) of an organism, includes the study of complex oligosaccharides present on glycoproteins. Considering the complexity of these oligosaccharides and their importance both clinically and scientifically, there exists a clear necessity for the development of suitable analytical platforms capable of characterizing the oligosaccharide structures. These glycomic platforms should include strategies for the separation, enrichment and the analysis of the oligosaccharides as well as methods of interpreting the resulting data. To achieve this goal a combination of different techniques from gel electrophoresis, chromatography, western and lectin blotting, nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) have been used.

MS, in particular, has become the analytical instrument of choice for the study of biological samples and is now the central component of any analytical glycomic platform⁹⁹⁻¹⁰¹. When MS is performed in combination with liquid chromatography (LC), it is particularly well suited to analysing a complex mixture of oligosaccharides released from a glycoprotein.

Liquid chromatography

With glycosylation rarely being a homogenous process, there is usually an assortment of different oligosaccharide structures including different isomeric forms of the same base sequence present on a glycoprotein. Using liquid chromatography to separate the individual oligosaccharides and their isomers prior to MS and MSⁿ analysis can improve the quality of data. LC can be performed online^{102,103} or offline^{104,105} or a combination of both^{106,107}. LC-MS of oligosaccharides is usually performed in capillary-flow (8-15 $\mu\text{l}/\text{minute}$) or nano-flow (0.3-2 $\mu\text{l}/\text{minute}$)¹⁰⁸. For LC-MS the choice of mobile phase will depend on its compatibility with the MS instrument. The choice of stationary phase will depend on whether analysis is performed on native or derivatized oligosaccharides. Although traditionally normal phase chromatography using either amine- or amide- type stationary phases have been commonly used for glycomic MS¹⁰⁹, we used reverse phase chromatography with Porous Graphitized Carbon (PGC) as our stationary phase. PGC is ideally suited for the separation of native oligosaccharides as it has good adsorption properties and offers high resolution separation of isomeric oligosaccharides. PGC is a highly homogenous crystalline structure of large graphitic sheets and is very well suited for glycomics as it can effectively separate different isomeric glycoforms¹¹⁰ in an oligosaccharide mixture prior to selection for fragmentation.

PGC has been used successfully to analyse *N*-linked¹¹¹⁻¹¹⁵, *O*-linked^{112,114,116-118} and sulfated oligosaccharides^{84,119,120}. Capillary flow is optimal for the analysis of native oligosaccharides in negative ion mode with PGC due to the poor ionization efficiency of electrospray ionisation (ESI) in high concentrations of water, however nanoflow is possible¹²¹.

There are fewer complications associated with nanoflow in positive ion mode. For derivatized oligosaccharides the type of derivatization will naturally determine the choice of chromatographic medium; for the separation of per-*O*-methylated oligosaccharides, C18 has become the most commonly used stationary phase for glycomics^{122,123}.

Mass spectrometry

Mass spectrometry is used routinely for the analysis of biological samples such as oligosaccharides. A mass spectrometer is an instrument that can detect and sort ions according to their mass to charge ratio (m/z). An MS can therefore be used to determine the mass of a particular analyte or analytes within a sample. Mass spectrometers consist of three principle components: 1) an ionization source 2) a mass analyser and 3) a detector (See Figure 5). The ionization source is used to ionize the analyte thus forming ions in a gaseous phase allowing them to enter the MS. The mass analyser separates the ions according to their m/z value and the ions are then sent to the detector to output their m/z value to the user.

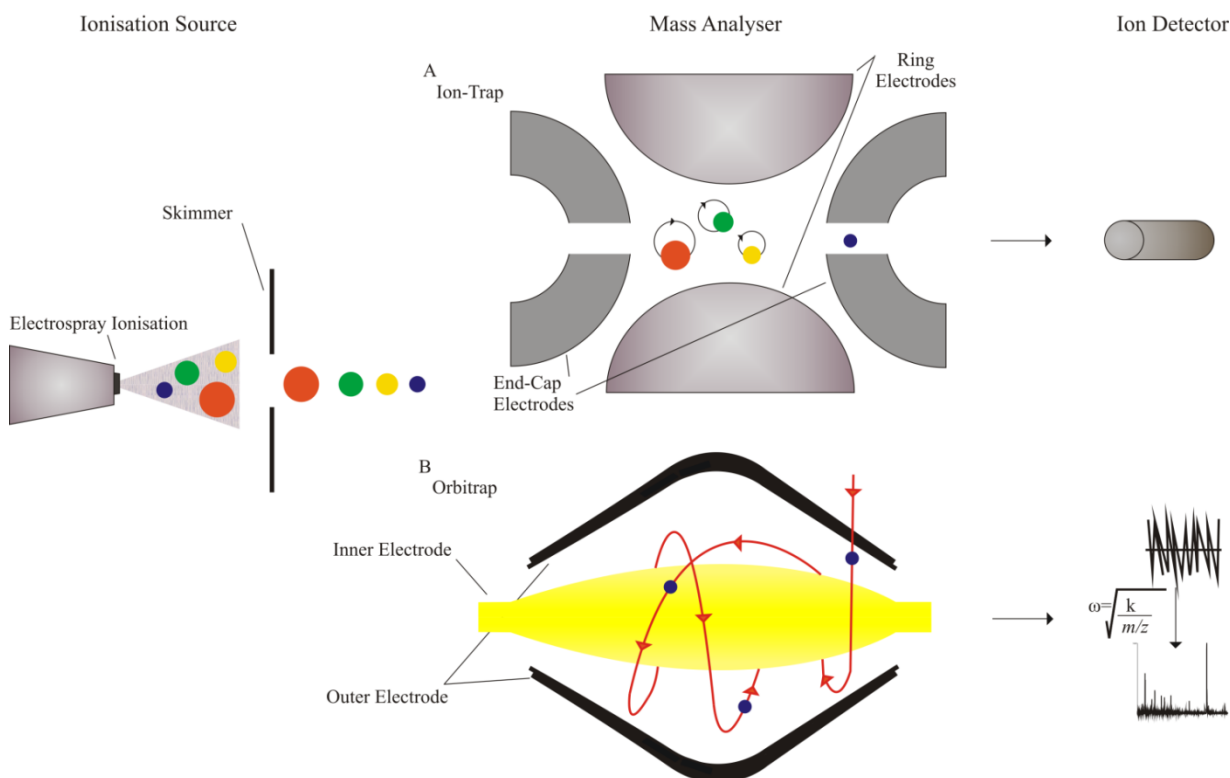


Figure 5: A mass spectrometer is made up of three constituent parts: An ionisation source, a mass analyser and an Ion detector. An electrospray ionisation (ESI) source is a ‘soft’ ionisation process ideal for the analyses of large biomolecules such as oligosaccharides. The skimmer aligns the ions prior to entering the mass analyser. A) In ion-trap mass analyser, an ion is confined by alternating charges on the end-cap electrodes and ring electrodes. The trapped ions rotate at a frequency that depends on their m/z value. A specific ion can be sent to an external detector which outputs its m/z value to the user, usually detected by an electron multiplier. B) An Orbitrap consists of a spindle shaped inner electrode and barrel shaped outer electrode. The ions are injected tangentially into the orbitrap and orbit the inner electrode, oscillating axially back and forth along the length of the inner electrode. The orbitrap acts as a detector as the ions oscillations is inversely proportional to the ions m/z value.

Electrospray Ionisation (ESI) has been used extensively for the study of large biomolecules, including oligosaccharides since its introduction by Fenn et al in 1989¹²⁴ due to its ability to ionize non-volatile molecules and all the MS data in this thesis was generated by ESI-MS. The general principle behind ESI involves the ionization of a solution containing the analyte passing through a capillary held under high voltage (between -3 and -5kV in negative ion mode)¹²⁵. This nebulizes the solution forming a gaseous mist. In capillary flow this process is usually aided by a drying gas such as nitrogen. The droplets are then reduced in size through evaporation and through a process of droplet subdivision known as ‘Coulomb explosion’¹²⁶. The ions, now in gaseous form, subsequently enter the MS to be analysed.

In matrix assisted laser desorption/ionisation or MALDI-MS, the analyte is ‘spotted’ onto a plate where it is co-crystallized with a matrix¹²⁷. The choice of matrix depends on many different factors and is crucial to successful ionization. Generally it is a low molecular weight acid that has good ultra violet (UV) or infra-red (IR) absorption properties; however basic matrices have also been used¹⁰¹. For glycomics, the choice of matrices has usually been the ones that were traditionally developed for proteomic analysis such as 2,5-dihydroxybenzoic acid (DHB). The trend towards using ionic liquid matrices (ILM’s), a mixture of a conventional acidic MALDI matrix such as DHB and an organic base has become a favoured choice for the analysis of sulfated oligosaccharides’ as they have been largely successful in limiting the disassociation of sulfates while simultaneously providing good MS sensitivity. Paper 1 of this thesis provides a more thorough introduction to MALDI-MS of sulfated oligosaccharides, including a description of the different ILM’s available.

The process by which the co-crystallized mixture is ionized involves a two-step process. After formation of the matrix-analyte crystallized spot, the desorption step is triggered by a UV or IR laser pulse which is applied to it. The matrix absorbs the laser energy causing ablation of the matrix-analyte mixture and forms a hot plume where the second step of ionization occurs. However the mechanism of this second step is still not fully understood and it is probable that more than one mechanism is involved in the ionization process. General belief is that through a process of protonation or deprotonation the matrix aids in the creation of charged species of the analyte¹²⁷.

Mass analyser and detector

After ionization, the charged ions are extracted and separated according to their mass to charge ratio (m/z) in the mass analyser. Different types of analyser are available such as quadrupole (QMS), triple quadrupole (QQQ), quadrupole ion trap (QIT), Fourier-transform ion cyclotron resonance (FTICR), time of flight (TOF) and Orbitrap and are suitable for different applications. All the results generated in this thesis were acquired using an ESI combined with either a linear QIT or an Orbitrap MS.

In a linear QIT a two dimensional (2D) RF field confines the ions radially and stopping potentials applied to end electrodes confines the ions axially¹²⁸. The ion trap captures ions under vacuum by a combination of magnetic and electrical fields. Apart from storing ions prior to detection, the QIT can also isolate particular ions based on their m/z value by regulating the RF field. The 'trapped' ions can then be subjected to fragmentation by collision induced dissociation (CID) for MS^n analysis. QIT's are considered low resolution mass spectrometers (Resolution is the ability of the MS instrument to distinguish two peaks of differing m/z value).

For high resolution MS an orbitrap, developed by the Russian physicist Alexander Makarov¹²⁹, can be used. The orbitrap is an inner spindle shaped electrode inside the outer barrel shaped electrode that creates an electric field. Ions are injected into the orbitrap where they 'orbit' the spindle-shaped electrode. The ions are trapped in this orbit as their electrostatic attraction to the inner electrode is counterbalanced by centrifugal forces as a result of their orbiting motion. The ions oscillate axially along the length of the inner electrode. It is possible to analyse the mass of the ions by sensing the ion oscillations, which are inversely proportional to the square root of their m/z value¹²⁹. For MALDI instruments, a Time of Flight (TOF) mass analyser is most often used. In a TOF, ions are accelerated under an electric field to have the same kinetic energy value, thus their velocity will be dependent on their m/z value. By measuring an ions time of flight within the analyser it is possible to determine its velocity and thus its m/z value. After separation the different ions are identified by a detector and their signal is converted to their respective m/z value.

Fragmentation of oligosaccharides

The base composition of an oligosaccharide can be assigned by interpretation of the MS data. This can provide information on the number of Hex, HexNAc, Fuc and sialic acid residues present on a particular oligosaccharide. We used MS data to assign the composition of *N*-linked oligosaccharides released from mammalian cell membranes in paper IV. For assignments of specific structural features, the oligosaccharide needs to be fragmented and the fragment ions analysed. This was the predominant method of analysing the *O*-linked oligosaccharides derived from the mucin MUC5AC in paper V.

The process of fragmenting an ion and analysing the fragment ions is known as MS^n , where 'n' is the number of times an ion has been sent for detection (therefore n-1 is the number of fragmentation events that have occurred). In MS^n , an ion that is referred to as the parent ion, is isolated for fragmentation. This fragmentation produces ions which are referred to as daughter ions or fragment ions. The daughter ions are detected by the mass spectrometer or isolated again for further fragmentation. By correctly interpreting the MS^n spectra it is possible to determine the composition of the parent ion.

For glycomic analysis fragmentation by collision induced dissociation (CID) has been used extensively for the purpose of identifying oligosaccharide structures^{116,130-133}. In CID the parent ion is isolated inside the ion trap which also contains an inert gas such as helium which stabilizes the parent ions. The parent ions are accelerated by increasing their RF voltage, thus increasing their kinetic energy, which causes the ions to collide with the resident inert gas. This collision will convert some of the kinetic energy into internal energy and result in bond breakage and fragmentation which produces daughter ions, which are then sent to the detector. CID is a particularly useful fragmentation technique for the analysis of oligosaccharides since it has predictable fragmentation properties and produces diagnostic fragment ions which can be used to assign structure. CID produces mainly glycosidic fragment ions which can be used to assign the primary sequence of the oligosaccharide. Cross ring X_j and A_i fragment ions are usually also present and can provide valuable information on the linkage position and branching within an oligosaccharide¹³⁴. We used cross ring fragments to identify the lacdiNAc epitope in paper V

CID has certain limitations, especially when analysing some sulfated oligosaccharides in negative ion mode¹³⁵ or fucosylated structures in positive ion mode^{136,137} as migration of these residues can occur which can complicate the interpretation of the MS^n spectra. Despite this, CID is a robust fragmentation pathway and was the predominant fragmentation technique used throughout this thesis.

High energy C-trap dissociation (HCD) is another fragmentation pathway available on the orbitrap mass spectrometers¹³⁸. This type of fragmentation was developed by Mathias Mann's group at Max Planck Institute of Biochemistry, Germany and utilizes a collision cell that is adjacent to the C-trap (an inverted C-shaped trap which is directly above the Orbitrap) within the orbitrap mass spectrometer to fragment ions¹³⁹. In HCD the ions are injected into the HCD cell at high energy and they collide with an inert gas which fragments the parent ion. HCD has a wider m/z range than CID; however a drawback to HCD is that it requires more ions than CID for fragmentation. This can increase the time required for fragmentation to occur, thus decreasing the number of MS^n spectra obtained via HCD. We found it difficult to obtain MS^n spectra that contained fragment ions from both glycosidic and cross ring cleavages from an oligosaccharide fragmented by HCD. Therefore deep structural analysis using HCD was particularly challenging. Despite this, HCD is viable alternative to CID for the analysis of sulfated oligosaccharides, as it avoids neutral loss and fragment rearrangement events during fragmentation that can occur during CID (Paper I & II)

Nomenclature for describing MSⁿ fragmentation

CID predominantly produces glycosidic and cross ring fragmentation. To describe these cleavage patterns, a nomenclature was developed by Domon and Costello¹⁴⁰. In Domon and Costello nomenclature, glycosidic fragments that include the reducing end of the oligosaccharides are designated B_i or C_j fragments whereas glycosidic fragments without the reducing end of the oligosaccharide are designated Z_j or Y_j fragments.

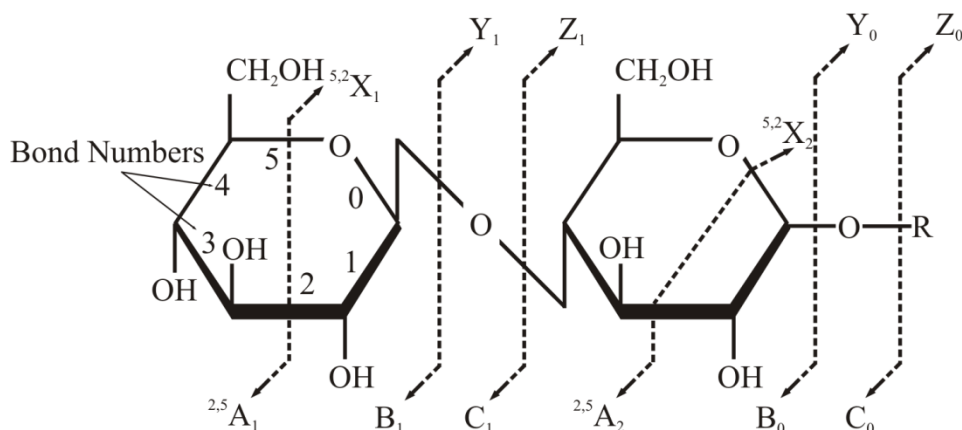


Figure 6: Domon and Costello fragmentation nomenclature (adapted from Domon & Costello 1988)

The *i* and *j* subscripts indicate the position of the subunit relative to the terminal ends of the oligosaccharide. Cross ring fragments that contain the reducing end of the oligosaccharides are labelled A_{*i*} fragments and those without the reducing end are designated X_{*j*} fragments. The specific cross ring cleavage is designated as ^{*k,l*}A_{*i*} or ^{*k,l*}X_{*j*}, where the *k* and *l* superscripts indicate the particular two carbon to carbon bonds that were cleaved across the cyclic sugar structure. For branching oligosaccharides; α , β , γ and δ are used to designate each branch, where α is designated to the largest branch, β to the next largest branch etc. (Figure 6)

LC-MS analysis of oligosaccharides

Characterisation of oligosaccharides by LC-MS can be very complicated due to the heterogeneity associated with glycosylation. Unlike amino acids chains where there is only a single linkage position possible, oligosaccharides can have as many as 4 branches and up to 6 different linkage positions with two anomeric forms (α or β). As a result, theoretically at least, there can be many different possibilities for structural assignment. This difficulty in structural elucidation is somewhat eased as in reality, where there is typically a limited number of glycosyltransferases available and present in the ER and Golgi apparatus during oligosaccharide biosynthesis. Therefore there are fewer linkage positions or isomeric possibilities to consider in a specific setting.

Furthermore, our knowledge of the glycosylation biosynthetic pathways allows us to predict certain glycosylation characteristics. For example, knowing that the core structure for all *N*-linked oligosaccharides is GlcNAc₂Man₃ or that lacNAc extension of *O*-linked oligosaccharides is either Type 1 (GlcNAcβ1-3Gal) or Type 2 (GlcNAcβ1-4Gal) can be useful information when assigning structures from MS data. The use of different software tools¹⁴¹⁻¹⁴³ as well as online databases¹⁴⁴⁻¹⁴⁷ can further ease structural assignment.

MS based glycomic analysis can be performed on either the native oligosaccharides or oligosaccharides that are derivatized prior to analysis. For MS of native oligosaccharides, analysis can be performed in either positive or negative ion mode. Negative ion mode is usually preferable especially if the oligosaccharides contain negatively charged residues such as sialic acids, hexuronic acids, sulfates or phosphates due to the ionization efficiency associated with their inherent negative charges. Analysis of native oligosaccharides in positive ion mode requires the introduction of mono-isotopic cations such as Na⁺ or K⁺. This can suppress the MS signal of the acidic species and furthermore interfere with the MS instrument by overloading it with charged salt ions.

For analysis in positive ion mode, derivatization of the oligosaccharide is therefore generally necessary to facilitate the ionization process. There are several different derivatization strategies available for the analysis of oligosaccharides¹⁴⁸. Of these, per-*O*-methylation¹⁴⁹⁻¹⁵¹ whereby the oxygen of a hydroxyl group (-OH) on the carbohydrate is substituted with a methyl group (-CH₃) is one of the most commonly used derivatization techniques. Despite the increased work-up procedure required for the analysis of per-*O*-methylated oligosaccharides, the MSⁿ spectra can include fragment ions that provide additional information on the linkage position of the monosaccharide residues¹⁴⁸. This can be particularly useful for determining linkage position of terminal sialic acids¹⁵² which, can be difficult with the analysis of native oligosaccharides without the use of sialidases.

Unlike proteomics where a variety of computational software tools, from search algorithms to online databases, are available and allow for automated processing of MS data, in glycomics the assignment of the oligosaccharide structures still requires manual interpretation of the MSⁿ spectra. This is due to the inherent complexity of oligosaccharides structures which can have multiple branches and different linkage position. Manual interpretation can thus be an arduous process, although it is somewhat easier than initially envisaged due to the reproducibility of the fragmentation patterns as well as the presence of particular diagnostic ions indicating specific cleavage events¹³⁴.

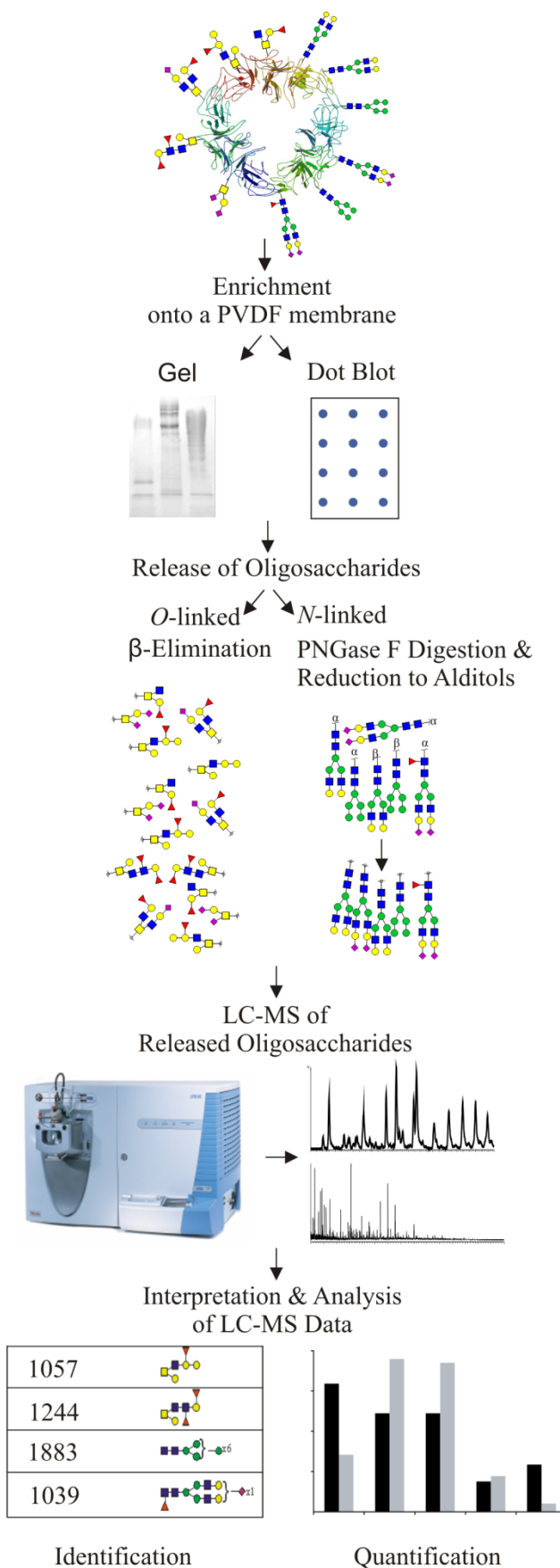


Figure 7: Typical workflow for the analysis of N- and O- linked oligosaccharides

In Figure 7 we present an overview of the workflow for our analytical glycomic platform for the analysis of *N*- and *O*- linked oligosaccharides. The glycoprotein is enriched prior to analysis which is usually necessary to remove contaminants that would otherwise reduce the quality of the LC-MS data. Hexose oligo- and poly-mers can be particularly troublesome contaminant that, if present in high concentrations, suppresses the MS signal of the oligosaccharides of interest. Our enrichment protocol typically involves immobilization of the glycoprotein to a Polyvinylidene fluoride (PVDF) membrane. If the glycoprotein of interest is present within a mixture of different proteins, SDS-PAGE (or SDS-agarose/polyacrylamide composite gels (SDS-AgPAGE) for larger glycoproteins such as mucins) is performed to separate the constituent parts of the glycoprotein mixture prior to the release of the oligosaccharides. For global analysis of a sample, direct transfer to a PVDF membrane via dot blotting is usually sufficient.

After enrichment, the oligosaccharides are cleaved from the protein. For *O*-linked glycosylation, the oligosaccharides are released by a β -elimination reaction with 0.5-1.0 M NaBH₄ in 50mM NaOH which is incubated at 50°C for 16 hours. The released oligosaccharides are then desalted by treatment with cation exchange resin to remove Na⁺ ions and repeatedly extracted with 1% acetic acid in methanol to remove the borate salts as borate esters. For *N*-linked glycosylation, the oligosaccharides are released by enzymatic digestion with 5 μ l of 0.5U/ μ l PNGase F and incubated overnight at 37°C. After PNGase F digestion, the released oligosaccharides are reduced to alditols in mild base in the same reaction that induces β -elimination of *O*-linked oligosaccharides. Post clean-up, the released oligosaccharides are analysed by LC-MS. The oligosaccharides are separated by gradient elution from a porous graphitized carbon (PGC) column and subsequently analysed by an ESI ion trap mass spectrometer in negative ion mode. Individual oligosaccharides are isolated for fragmentation by CID and their structures is assigned by manual interpretation of MS and MSⁿ spectra. Further analysis of the LC-MS data can be used to derive quantitative information which can provide a deeper insight into the glycosylation of the glycoprotein.

Alternative methodologies for the characterization of the glycome

Our glycomic platform is well suited to the characterization of native *N*- and *O*- linked oligosaccharides. However other methodologies have been adopted for the glycomic analysis. Matrix Assisted Laser/Desorption Ionisation (MALDI) is another soft ionization technique and has been used extensively for the analysis of oligosaccharides¹⁵³⁻¹⁵⁸. However, MALDI MS is incompatible with an online-LC setup. For samples requiring separation prior to analysis, such as the *N*-linked oligosaccharides we released from the cellular membranes (Paper IV) or the *O*- linked oligosaccharides released from mucins (Paper V) where a high degree of heterogeneity exists and presence of different isomeric structures is common, MALDI is less suitable.

Our approach has been to analyse the released oligosaccharide in its native form or by reduction to alditols. However as an alternative to this, derivatization of the oligosaccharide prior to LC-MS analysis is often performed. Of the derivatization techniques available, per-*O*-methylation is one of the most commonly used in glycomic MS¹⁴⁸.

Furthermore, instead of analysing the released oligosaccharide it is also possible to perform glycoproteomic analysis whereby the oligosaccharides are still attached to the peptide¹⁵⁹⁻¹⁶¹.

Per-*O*-methylation

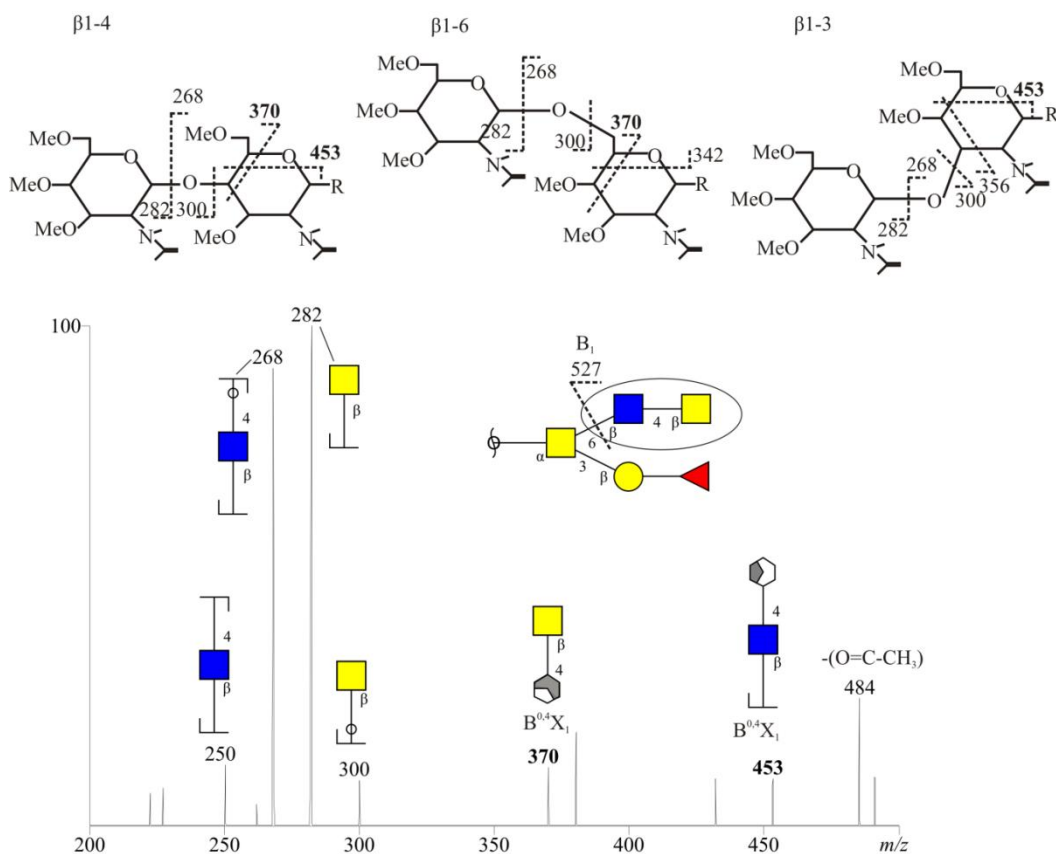


Figure 8: MSⁿ fragmentation of a per-*O*-methylated *O*-link Fucα1-2Galβ1-3(GlcNAcβ1-4GlcNAcβ1-6)GalNAc-ol with an [M-H]⁻ *m/z* 936. The MS² of this structure contained a fragment with an *m/z* 527 corresponding to B₁ ion of GalNAcβ1-4GlcNAcβ1-. This fragment ion was isolated for MS³ fragmentation and contained two fragment ions, *m/z* 370 and *m/z* 453, that were specific to a 1-4 linkage for the terminal GalNAc residue. This provided unambiguous assignment of the linkage position.

Per-*O*-Methylation involves the substitution of a methyl functional group for all the hydroxyl functional groups on the oligosaccharides. Per-*O*-methylation has been widely used for the analysis of oligosaccharides^{103,149,162}. The per-*O*-methylation protocol increases the preparation time of the oligosaccharide post release. Furthermore the traditional per-*O*-methylation techniques^{150,151} have shown to be problematic for the analysis of sulfated oligosaccharide due to the labile nature of the sulfate group which makes the post derivatization work-up procedure difficult¹²⁰, however recent advances in the development of reversed phase clean-up procedures for sulfated oligosaccharides have overcome some of these challenges^{149,163}.

Per-*O*-methylation converts the sialic acids to neutral esters. Although this sacrifices the increased ionization efficiency of the acidic species in negative ion mode, it improves the consistency of ionisation in positive ion mode of both neutral and acidic residues which can aid in the quantification of oligosaccharides¹⁶⁴. MS analysis of per-*O*-methylated oligosaccharides can provide more informative spectra and provides unambiguous assignment of linkage position¹⁶⁵ and this is exemplified in Figure 8 where we have per-*O*-methylated the lacdiNAc containing m/z 936 structure present on HGM-2 (see Paper V). We can see that the MS³ of the B₁ ion produces two fragment ions, m/z 370 and m/z 453 that are specific to a 1-4 linkage for the terminal GalNAc residue.

Glycoproteomics

Our current methodology of cleaving the oligosaccharides from the protein prior to the analysis by LC-MS means that information on the protein portion of the glycoprotein is lost, including the particular site of glycosylation. Traditional proteomic analysis will deduce minimal information on the carbohydrate content of a glycoprotein. In glycoproteomics, intact glycopeptides are analysed and this can provide information on both the protein and oligosaccharide¹⁶⁶. Enrichment of glycopeptides is a prerequisite to any glycoproteomics experiment as generally the glycopeptides are the minor component of the peptide mix. Glycoproteomics is particularly well suited to the analysis of *N*-linked glycosylation where the presence of the consensus sequence means that it is possible to predict the glycosylation site and this information can be used during experimental design. Glycoproteomic analysis of *O*-linked glycosylation has been more limited. The lack of a consensus sequence means it is impossible to predict the site of glycosylation by the primary amino acid sequence^{54,167}. Furthermore, *O*-linked glycosylation is generally more heterogeneous than *N*-linked glycosylation and this adds a level of complexity to the resultant spectra. Glycoproteomics of the *O*-glycosylated glycoproteins is limited. This is because the oligosaccharides are usually found in high density glycosylated area (mucin domains) that limits the effectiveness of various proteases. The current *O*-linked glycopeptides detected are almost exclusively limited to less glycosylated areas

As with glycomics, glycoproteomics generally relies upon manual interpretation of the MSⁿ spectra for structural assignment. In glycoproteomics, acquiring the MSⁿ spectra of glycopeptides requires two fragmentation events; one for the analysis of the oligosaccharide sequence (usually CID or other low energy collision that preferentially fragment the oligosaccharide part) and one for the analysis of the peptide (subsequent CID of unglycosylated or low glycosylated fragments, or alternatively ECD or ETD of intact glycopeptides). The analysis of the data increases the time required to annotate and assign structure. The glycoproteome is considerably more diverse than the glycome, thus global analysis of the glycoproteome is generally unfeasible¹⁶⁸. As such, a proteomic and glycomic analysis continues to be a powerful analytical approach for the characterization of a glycoprotein. However, glycoproteomics is becoming increasingly important¹⁶⁹ and a comprehensive understanding of glycosylation requires knowledge of the site-specific location of the oligosaccharides.

1.3 MS analysis of oligosaccharides

MS analysis of *N*- and *O*- linked oligosaccharides

Analysis of *N*-linked oligosaccharides is typically performed after enzymatic cleavage of the oligosaccharides by a Peptide: *N*-glycosidase (PNGase), either from the bacterium *Flavobacterium meningosepticum* (PNGase F¹⁷⁰) or from almond (PNGase A¹⁷¹) which releases the oligosaccharides with their reducing end intact. MS analysis of sialylated *N*-linked oligosaccharides can be difficult due to the labile nature of sialic acids which causes the sialic acid residue to readily disassociate during MSⁿ fragmentation. As a result, the MSⁿ spectra shows very limited fragmentation apart from the dissociation of the sialic acid, thus providing little information about the rest of the oligosaccharide. It may therefore be necessary to remove sialic acids prior to MS analysis in order to obtain information-rich MSⁿ spectra of the rest of the oligosaccharide. Per-*O*-methylation has been shown to stabilize the sialic acid residues by converting the carboxylic acid to neutral esters, resulting in more information rich MSⁿ spectra of sialylated structures¹⁴⁹. MS of per-*O*-methylated *N*-linked oligosaccharides can provide additional information due to the increased number of cross ring fragmentations, although the spectrum can also be more complex and this can make interpretation of the data more challenging, especially for the novice glyco biologist.

Typically reductive β -elimination under alkaline conditions is used for the isolation of *O*-linked oligosaccharides. The oligosaccharides are released from this process as reduced alditols. Although this means that the oligosaccharide cannot be labelled, for example with 2-aminobenzamide (2-AB), it is the preferred method of release as it avoids destructive peeling reactions that can occur with other release methods such as mild hydrazinolysis¹⁷². It can be advantageous to separate the neutral oligosaccharides from acidic oligosaccharides prior to MS analysis. This is especially important for some mucins which can be heavily sialylated. This is particularly common in samples derived cancerous tissue where the increased sialylation can be used as a biomarker for the presence of a tumor.

The separation of the neutral and acidic species can aid MS analysis of non-derivatized oligosaccharides in negative ion mode as it avoids any preferential ionisation of the charged acidic oligosaccharides which could result in signal suppression of the neutral structures. Anion exchange chromatography is a good choice for separating sialylated and sulfated structures from their neutral counterparts. For online LC-MS of *O*-linked oligosaccharides from mucins, online chromatographic separation should be particularly well suited to the separation of isomers as it is common to find the several oligosaccharides with the same m/z value on mucins.

MS analyses of sulfated oligosaccharides

MS analysis of sulfated oligosaccharides can be problematic due to the labile nature of the sulfate residue resulting in the sulfate disassociating from the oligosaccharide during ionisation. There are further complications when analysing *N*- and *O*-linked oligosaccharides due to the lower abundance of sulfated species, thus they may not be detected during MS. Even in samples where the major component are sulfated oligosaccharides such as GAG's, the MSⁿ spectra from multiply sulfated residues can be very complex which can be difficult to correctly interpret. For the detection of low abundant sulfated *N*- and *O*- linked oligosaccharides, an enrichment step should be initiated prior to MS analysis. Apart from the standard 1D and 2D SDS-PAGE approach for enrichment of glycoproteins, various other enrichment strategies are available. Amongst these alternative enrichment steps are for example enrichment of acidic oligosaccharides (including the sulfated oligosaccharides) from their neutral counterparts by anion exchange¹⁷³. To specifically target sulfated oligosaccharides, removal of the sialic acids via treatment with sialidase may be performed prior to anion exchange. Other enrichment strategies designed to specifically target sulfated oligosaccharide include enrichment by strong cation exchange on a gel permeation media¹⁷⁴ or by a process known as the 'sulfate emerging procedure' developed by Toyoda et al¹⁷⁵.

Per-*O*-methylation of the sulfated oligosaccharides has, until recently, been largely unsuccessful as the labile sulfate group is lost during the workup procedure when using the traditional per-*O*-methylation protocols^{150,151}. To avoid this loss of sulfate groups from occurring, an alternative approach to the per-*O*-methylation procedure using solid phase clean-up with C18 has been developed^{149,163}. Furthermore the per-*O*-methylation converts the charged sialic acid residues to neutral per-*O*-methylated esters. This could result in the sulfated species being the sole acidic oligosaccharides present in a mixture and it would therefore be possible to include an enrichment step designed specifically for the enrichment of per-*O*-methylated sulfated oligosaccharides using anion exchange chromatography¹²⁰. A more in-depth review of MS analysis of sulfate *N*- and *O*- linked oligosaccharides is presented in paper I of this thesis.

2 Aims of Thesis

The general aim of this thesis was to analyse complex mixtures of *N*- and *O*- linked oligosaccharides with a particular interest in identifying potential biologically relevant glyco-epitopes. More specifically we were interested in:

- Improving the acquisition of MSⁿ spectra of sulfated oligosaccharides by investigating the migration of sulfates during CID fragmentation and identifying methods to reduce the migration from occurring.
- Developing a protocol for the enrichment of membrane associated glycoproteins for glycomic analysis.
- Refining the process of displaying MS data by incorporating different statistical methodologies.
- The characterisation of glyco-epitopes which have been identified as potentially being biologically relevant.

3 Results and Discussion

Unlike the synthesis of proteins, glycosylation is a non-template driven process. The final composition of an oligosaccharide is instead determined by the actions of various glycosyltransferases and glycosidases as well as transferases for substitution, *e.g.* sulfotransferases, as well as the availability of donor monosaccharides within the Golgi apparatus. These biosynthetic reactions can be affected by intra- and extra- cellular stimuli and may not always go to completion¹⁷⁶. The glycome is therefore a complex array of different oligosaccharides. This inherent heterologous nature can frustrate the analytical process as the mass spectrometer is unable to differentiate the various glycoforms that are typically present in an oligosaccharide mixture. The data in this thesis suggest that it is preferable if the various glycoforms are separated prior to MS analysis. This is particularly relevant if the oligosaccharide structure is going to be characterized by fragmentation. Several methods of separating the glycoforms have been used for glycomic analysis, from offline chromatographic separation of the oligosaccharides^{104,107,177} to separation by gel electrophoresis^{154,178,179} prior to MS analysis. Online LC-MS provides an elegant and effective alternative method of separating and analysing released oligosaccharides and is ideally suited to the analysing heterogeneous mixtures of *N*- and/or *O*- linked oligosaccharides.

For LC-MS, PGC is an effective stationary phase suitable for glycomic analysis of native oligosaccharides. The resolving power of PGC is clearly displayed in Figure 9 where oligosaccharide isomers, all with a $[M-H]^-$ m/z of 733, isolated from human gastric mucin is present as five different glycoforms. Using PGC, the five isomers were separated into distinct chromatographic peaks allowing for unambiguous assignment of their respective structures by MS.

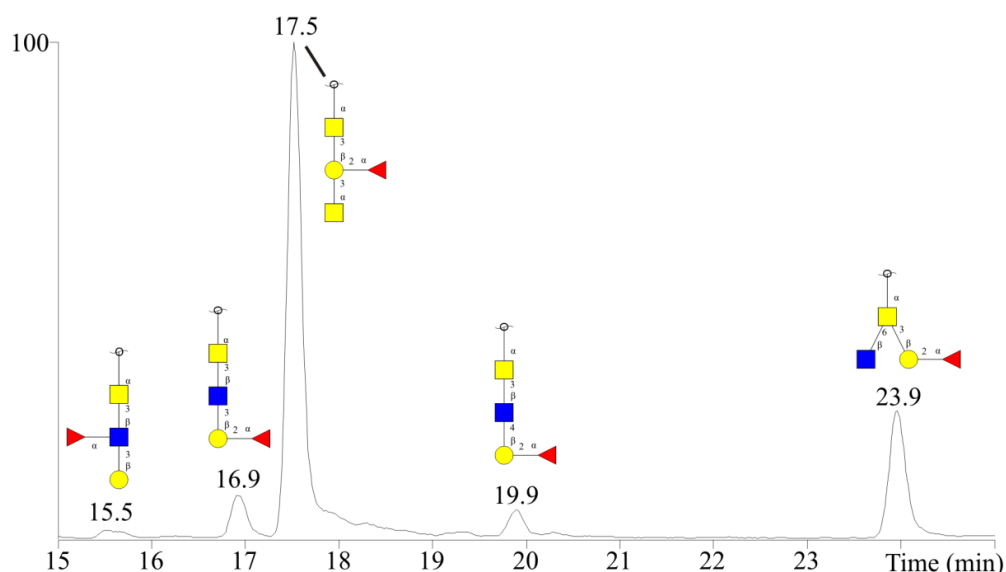


Figure 9: An extracted ion chromatogram showing five isomeric mucin type oligosaccharides all with an $[M-H]^-$ of m/z 733 separated by PGC. The five isomers were separated into distinct chromatogram peaks allowing for easier assignment of the oligosaccharide structures from the MS^n spectra.

The process of manually interpreting MS and MSⁿ spectra is benefiting from an increasing availability of software tools and databases. Glycoworkbench can assist in the manual interpretation of MS spectra¹⁴² and Glycomod can determine the possible composition of an oligosaccharides based on experimentally determined masses¹⁴¹ and both were used routinely when assigning oligosaccharides structures from MS data. Online databases of oligosaccharides, such as the UniCarb-DB¹⁴⁴, can be used to identify particular oligosaccharides within a sample. The reproducibility of CID fragmentation makes it possible to directly compare the MSⁿ fragmentation pattern of unknown oligosaccharide to the MSⁿ fragmentation of a structure that has previously been identified or to oligosaccharide standards where the spectra are accessible from glycomic databases. This approach was also invaluable when analysing the *O*-linked oligosaccharides released from gastric MUC5AC in Paper V¹³³.

3.1 Migration of sulfates

CID fragmentation has become a standard fragmentation pathway for the analysis of oligosaccharide. However, rearrangements events whereby residues within the oligosaccharide migrate during CID fragmentation can occur. These rearrangements are problematic as they complicate the interpretation of the spectra and mis-assignment of the oligosaccharide structure can occur. The CID induced migration of fucose from one monosaccharide to a different monosaccharide during CID fragmentation in positive ion mode is a well-known migration event^{136,137}. Fragmentation of oligosaccharides in negative ion mode provides information rich MSⁿ spectra without the migration of fucose occurring. Furthermore negative ion mode MS of oligosaccharides is particularly favourable for the analysis of acidic species as the inherent negative charge improves the ionisation efficiency. However, negative ion mode CID is not without its drawbacks.

In Paper I we introduced the phenomenon of the migrating sulfate whereby a sulfate residue is transferred from one monosaccharide to another during negative ion mode CID fragmentation. As a result of this migration occurring, the interpretation of the MSⁿ spectra of sulfated oligosaccharides proved problematic and provided non-conclusive information on the location of the sulfate residues. The sulfate migration was initially observed during the analysis of *O*-linked oligosaccharides released from ascites glycoproteins. A sulfated sialylated *O*-linked pentasaccharide was present as the [M-H]⁻, [M+Na-2H]⁻ and [M-2H]²⁻ ion. During CID fragmentation of the [M-H]⁻ ion, we observed the presence of the diagnostic fragment ion *m/z* 370. This particular ion indicated that the sulfate residue had migrated to the terminal sialic acid. We also noted that migration was not detected in the MSⁿ spectra of the [M+Na-2H]⁻ or the [M-2H]²⁻ ions (Paper I, Figure 6). These initial results suggested that migration was being promoted by the presence of a mobile proton, the presence of which can influence the quality of CID spectra¹⁸⁰.

In Paper II we investigated sulfate migration further, with the aim of identifying the process by which the migration occurs. We fragmented various sulfated oligosaccharides by CID in negative ion mode and measured the degree of sulfate migration (Paper II, Figure 2).

The results showed that the amount of sulfate migration varied greatly between the different sulfated structures analysed and we could see the level of sulfate migration increased dramatically when a terminal sialic acid was present. The increased level of migration occurring on $[M-H]^-$ of the sialylated-sulfated structures supported our earlier findings that migration was being promoted by the presence of a mobile proton as we could show that sulfate migration does not occur when the mobile proton was unavailable such as in the MS^n fragmentation of the $[M+Na-2H]^-$ (Paper II, Figure 5) or the $[M-2H]^{2-}$ (Paper II, Figure 7). During this initial analysis, we also observed that sulfate migration was increased when an oligosaccharide was reduced to the more flexible alditol compared to the more rigid pyranose ring containing aldose version of the same oligosaccharide structure. We also could show that sulfate migration was increased when the sulfate was present on the more flexible 6th carbon of CS-6 which is outside of the pyranose ring rather than on the 4th carbon of CS-4 which is within the ring when we fragmented chondroitin sulfate disaccharides (Paper II, Figure 4).

The presence of a mobile proton alone would not explain this increased level of sulfate migration and we therefore wanted to identify the reason for this difference. The results suggested that migration was also promoted by the increased steric availability of the donor oligosaccharide. To explore the relationship between sulfate migration and steric nature of the sulfate oligosaccharides further, we isolated the acceptor monosaccharides for MS^3 analysis to see where on the monosaccharide the sulfate was migrating to. The MS^3 spectra showed that the sulfate was migrating to the freely rotating 6th carbon on Hex and HexNAc residues as well as the flexible glycerol side chain of sialic acid (Paper II, Figure 6). With these results we could conclude that the generation of the mobile proton is a prerequisite for migration to occur and that increasing flexibility of both the donor and acceptor monosaccharides lead to increased levels of sulfate migration (see Figure 10).

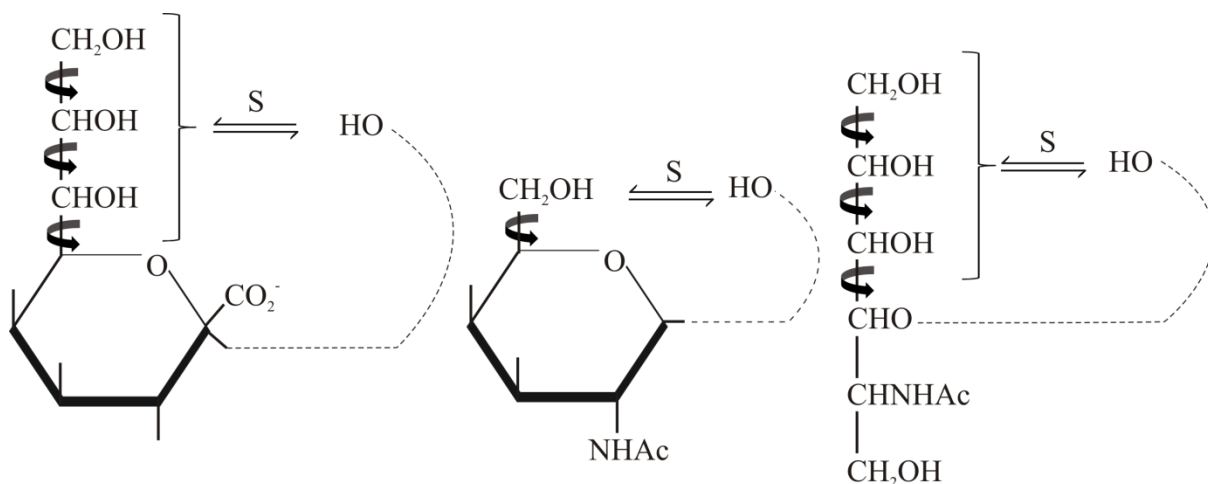


Figure 10: Summary of pathways for migration of sulfate in CID. Proposed migration of sulfate to and from areas of high degree of rotational freedom in oligosaccharides containing sialic acid, hexoses and alditols with a free rotational C-6 promoted by the generation of a mobile proton.

Alternative fragmentation pathways: minimising sulfate migration

Having identified that the presence of a mobile proton as well as a high degree of conformation freedom within the oligosaccharides induced the migration of sulfate, we broached the question of how to avoid sulfate migration from occurring. Our initial approach to limit sulfate migration was to reduce the generation of the mobile proton by preferentially choosing the $[M+Na-2H]^-$ or the $[M-2H]^{2-}$ ions of sulfated oligosaccharide for MS^n fragmentation. However, we considered neither of these strategies satisfactory. The inclusion of a cation such as Na^+ into the sample suppresses signal intensity in negative ion mode. With the possibility that the sulfated oligosaccharides are already a minor portion this was considered undesirable. We also found it difficult to limit the amount of mobile protons by increasing the charge state of the parent ion. This was especially true of oligosaccharides where the sulfate is the sole charge carrying residue. Our efforts to reduce migration by changing the CID parameters during fragmentation were also unsuccessful.

With our limited success at reducing sulfate migration with CID, we explored the possibility of using alternative fragmentation pathways. Higher Energy C-trap Dissociation (HCD)¹³⁹, a fragmentation pathway available on Orbitrap MS instruments, had previously been used to analyse sulfated peptides where no neutral loss of the sulfate group during fragmentation was detected¹⁸¹. We therefore investigated HCD as a possible alternative to CID for the fragmentation of sulfated oligosaccharides. The MS^n spectra of sulfated and sulfated-sialylated oligosaccharides fragmented by HCD provided informative MS^n spectra of the $[M-H]^-$ ions without inducing migration of the sulfate residue (Paper II, Figure 8)

The initial results suggest that HCD could be a suitable alternative to CID for the analysis of sulfated oligosaccharide. That said, HCD fragmentation has a longer acquisition time than CID, thus fewer MS^n spectra can be accumulated throughout an LC-MS run. CID is such a well-established fragmentation pathway for glycomics that it is likely to remain a prominent fragmentation pathway. The use of alternative fragmentation pathways such as HCD would be complimentary to CID for the fragmentation of oligosaccharides. It is possible to perform CID and HCD at the same time which could be an efficient method of accumulating good MS^n data for sulfated and non-sulfated oligosaccharides. The availability of HCD as a fragmentation pathway that avoids migration of sulfates (and neutral loss of the sulfate residues) is clearly beneficial to MS based glycomics.

3.2 LC-MS of *N*- and *O*- linked oligosaccharides

***N*-linked oligosaccharide characterization: A semi quantitative approach**

In Paper III we describe an efficient protocol for the global enrichment of membrane bound proteins from lysed cells by a process of carbonate extraction¹⁸² and high molecular mass membrane bound glycoproteins from milk by process of ultracentrifugation. The protocol also describes the workflow for the preparation of the samples for glycomic analysis.

This includes the transfer of the enriched glycoproteins to PVDF membranes via a SDS-AgPAGE gel or by dot blotting, the release and clean-up of the *N*- and *O*- linked oligosaccharides from the glycoproteins as well as preparation of a PGC chromatography column and subsequent analysis of the oligosaccharides by LC-MS.

To highlight the effectiveness of our global enrichment protocol, in Paper IV we analysed the *N*-linked oligosaccharides of two enriched membrane samples shown to differ in their degree of sialylation. Our analysis of the LC-MS data of the released *N*-linked oligosaccharides showed that the membranes had very different glycosylation profiles with the high sialic acid (HSA) membrane containing oligosaccharides mainly with one or two sialic acids while the low sialic acid (LSA) membrane contained mainly unsialylated oligosaccharides (Paper IV, Figure 1). Of the total of the 41 oligosaccharide compositions that we identified, 21 compositions were present on both HSA and LSA, with a further 20 composition detected on LSA only. Our initial analysis of the LC-MS data was based on the assignment of the composition based on the m/z value of the oligosaccharide, thus limiting our assignment to the number Hexoses (Hex), *N*-acetylhexosamines (HexNAc), deoxyhexoses (i.e. considered to be Fucose, Fuc) and sialic acids in a particular oligosaccharide.

We were not satisfied with limiting our characterisation to such a rudimentary assignment of oligosaccharide structure. To better describe the membrane oligosaccharides, we applied our knowledge of the *N*-linked biosynthetic pathway to the data generated by LC-MS. By doing so we could assign specific monosaccharide residues to the compositions that were identified. By applying this process, we were able to assign high-mannose and complex structures on both membranes analysed. We could detect high-mannose structures with up to 9 mannose residues on both LSA and HSA. Of the complex structures that we identified, 16 were detected on both LSA and HSA and consisted of sialylated and unsialylated structure with and without core fucosylation. The remaining complex structures were detected on LSA only and were all unsialylated, extended with up to 12 lacNAc residues (Paper IV, Table 1).

Generation of semi-quantitative data

This initial approach to representing the data provided a general overview of the glycosylation profile, showing the presence or absence of specific oligosaccharide structures. However, with glycosylation having been described as less of an on-off ‘digital’ and more of a subtle ‘analogue’ process¹⁸³, there is a requirement to highlight changes in the glycosylation profile, rather than just the presence or absence of a particular structure. We believed that the lack of quantitative data limited the amount of biologically relevant information that we could derive from the data as presented in Table 1 of Paper IV. We therefore generated semi-quantitative results from the LC-MS data and this showed that there was significant variation in the relative abundance of the different oligosaccharides. The semi-quantitative data that we had compiled was initially processed as a list of structures on the x-axis and their relative abundance as a percentage of the overall oligosaccharide content on the y-axis (Paper IV, Figure 2).

However we considered this approach of displaying the data was not optimal as it provided us with little information on the overall glycosylation and it did not make any connection between the identified structures and their biological function. To display the data within a global context, we decided to display the semi-quantitative data using mass spectrometry average composition (MSAC)¹⁸⁴ which reduces each sample to their monosaccharide composition. The MSAC data was ideal at unambiguously showing the differences in the level of sialylation as well as providing details on the levels of Gal and GlcNAc, which gives an indication on the amount of LacNAc extensions (Paper IV, Figure 3A).

We were also interested in narrowing our focus and sorting the oligosaccharides by particular glycosylation features and thus garner additional information about the membrane glycosylation. By looking at the differences between the relative abundance of high-mannose to complex structures we could see there was a similar ratio for both membranes, with the majority of the structures present being complex (Paper IV, Figure 3B). When we compared the sialylated and unsialylated complex structures, we could show that the sialylated structures were the major component of HSA and the unsialylated structures were the major component of LSA (Paper IV, Figure 3C). The results generated through the analysis of membrane associated oligosaccharides show that our glycomic LC-MS platform is well suited for the analysis of *N*-linked oligosaccharides. It was apparent from our initial analysis, that there is no ‘one size fits all’ approach to presenting glycomic data and that a combination of different methodologies may be required depending on the type of information that needs to be displayed and that the use of semi-quantitative data was beneficial at reflecting the ‘analogue’ nature of the glycome.

In Paper IV we have used MSAC to display the glycosylation within a global context while also narrowing our focus and by comparing specific glycosylation features such as high-mannose v’s complex and sialylated v’s unsialylated. However, these were not the only methods of displaying information available. While examining *O*-linked oligosaccharides released from human gastric MUC5AC (Paper V), we considered using heat plots to cluster samples with similar glycosylation profiles. By clustering samples with similar attributes, it may be possible to identify and correlate specific features of their glycosylation profile with their respective interaction with *H. pylori*. While this was certainly a promising method of analysing results and has been used previously to display glycomic data¹⁸⁴, the limited number of samples available ultimately meant that this was not applicable. However, it is evident that the manner in which the data is displayed can be as important as the data itself.

Characterization of MUC5AC derived *O*-linked oligosaccharides

We have also used LC-MS to characterise the mucin type *O*-linked oligosaccharides of MUC5AC. In Paper V, *O*-linked oligosaccharides derived from human gastric mucin (HGM) were analysed by LC-MS. We released *O*-linked oligosaccharides from the protein backbone by reductive β -elimination and subsequently analysed the released oligosaccharides by LC-MS. We assigned the oligosaccharide structures by interpretation of the MSⁿ spectra in conjunction with LC retention time to differentiate the various isomeric glycoforms.

The oligosaccharides were isolated for fragmentation by CID which yielded mostly Z_i and Y_i fragment ions which enabled us to assign the primary sequence and the presence of particular fragmentation patterns or diagnostic ions present in the MS^n spectra provided the additional residue and linkage information¹³⁴. In addition to the manual interpretation of MS^n spectra, our identification of the structures was aided by comparing the MS^n spectra of oligosaccharides present on HGM with the MS^n spectra from previously identified oligosaccharides using the UniCarb-DB database.

For specific characterization of HGM *O*-linked glycosylation, we characterized the *O*-linked oligosaccharides from MUC5AC isolated from the normal tissue of a patient with gastric cancer (HGM-1) and isolated from a healthy individual (HGM-2). The LC-MS showed that HGM-1 and HGM-2 had very different glycosylation profiles (Paper V, Figure 2). LC-MS of both samples revealed that HGM-1 had a very heterogeneous glycosylation profile with a high abundance of different glycoforms compared to HGM-2, which was more homogenous and dominated by a small number of neutral oligosaccharides. In total, we identified 47 individual structures on HGM-1 and 12 individual structures HGM-2. We were able to utilise the presence of particular diagnostic ions¹³⁴ to show that core 1, 2 and 3 structures were present on both MUC5AC samples and that core 4 structures were also present on HGM-1.

We identified oligosaccharides which were terminating with blood group H (Fuc α 1-2Gal β 1-). This termination can only occur through the activity of the secretor-gene (α 2 fucosyltransferase) which was present on both HGM-1 and HGM-2, which indicated that both samples were isolated from secretor individuals. Initially, we identified this particular terminal structure on an oligosaccharide with an $[M-H]^-$ m/z of 530 (Fuc α 1-2Gal β 1-3GalNAc-ol), however further analysis revealed other blood group H containing oligosaccharides which confirmed that HGM-1 and HGM-2 were isolated from secretor individuals.

Since both individuals were secretors, it is possible to determine their blood group status by identifying the presence of oligosaccharides containing specific blood group epitopes. We could identify oligosaccharides containing blood group A (GalNAc α 1-3(Fuc α 1-2)Gal β 1-) and blood group B (Gal α 1-3(Fuc α 1-2)Gal β 1-) (as well as the blood group H) epitopes on HGM-1 indicating that the individual has blood group AB, whereas we only identified the blood group H structures on HGM-2. Assignment of Type 1 and Type 2 extension relies on the presence of the cross ring ^{0,2}A_i fragment ion of the C-6 GlcNAc. This particular fragment ion is consistently produced by the fragmentation of Type 2 (Gal β 1-4GlcNAc1-) extended oligosaccharides. The absence of this fragment ion indicated the presence of a Type 1 (Gal β 1-3GlcNAc1-) extension. Due to the low mass cut off in the ion trap, predicted ^{0,2}A_i fragment ions were sometimes out of the observed mass range, and we could therefore not assign Type 1 or Type 2 extensions. In addition to complex extension, we also detected shorter oligosaccharides such as the cancer associated T (Gal β 1-3GalNAc-ol $[M-H]^-$ m/z 384), Sialyl Tn (NeuAc α 2-6GalNAc-ol $[M-H]^-$ m/z 513) and Sialyl-T (NeuAc α 2-6(Gal β 1-3)GalNAc-ol $[M-H]^-$ m/z 675) antigens and were present only on the tumour associated MUC5AC of HGM-1 and absent on the healthy MUC5AC of HGM-2. This observation needs to be extended by the analysis of additional samples.

MSAC analysis of mucin *O*-linked oligosaccharides

As previously discussed, in Paper IV we investigate different methods of presenting glycomic LC-MS data. We found that generating semi-quantitative data was beneficial as it could highlight differences in the glycosylation profile of glycoproteins. MSAC analysis, in particular, was a useful statistical tool for the global analysis of oligosaccharides. We have therefore also applied MSAC analysis to the *O*-linked oligosaccharides from gastric MUC5AC samples.

We characterised the *O*-linked oligosaccharides from a further 8 individuals (1 healthy sample, 3 tumor and 3 non-tumor samples from cancer patients) and have combined the LC-MS data with the LC-MS data from the two HGM samples from Paper V to generate MSAC results. Our MSAC analysis of these samples (Figure 11) shows that both the healthy and the non-tumor tissues share a similar monosaccharide content, with low amounts of sialic acid and higher amounts of fucose compared to the tumor MUC5AC. One aspect of MSAC analysis of mucin *O*-linked oligosaccharides is that the secretor status of the individual will affect their monosaccharide content. In our analysis, the non-tumor and tumor samples were all derived from secretor individuals, however the healthy samples were derived from one secretor and one non-secretor individual. The large error bar for Fuc in the healthy samples is a result of their secretor status. We see in the LC-MS data of both of these individuals (data not shown) that they both have a similar homogenous glycosylation profile. However, on the secretor individual (HGM-2 from paper V) many of the prominent oligosaccharides are terminated by the blood group H antigen (thus will contain one or more Fuc residues) whereas the oligosaccharide from the non-secretor were largely unfucosylated and this difference accounts for the large error in the level of Fuc on the healthy individuals in the MSAC data.

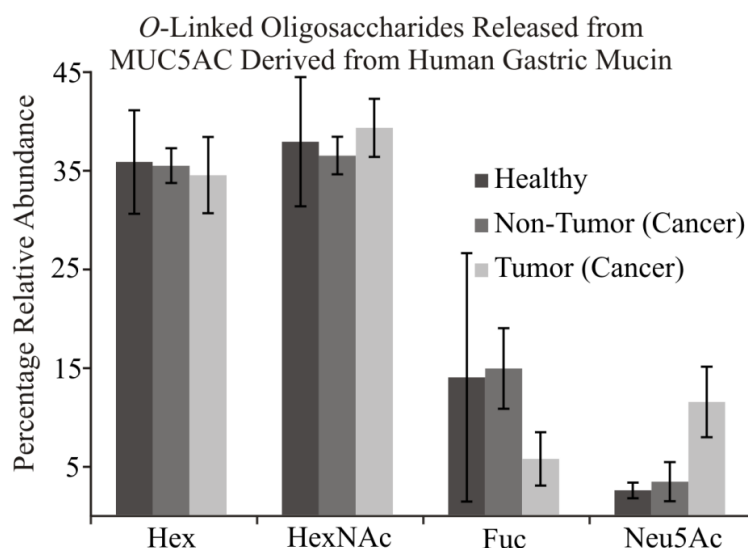


Figure 11: MSAC analysis of *O*-linked oligosaccharide released from MUC5AC derived from human gastric mucin. MUC5AC samples were isolated from patients without cancer (healthy), from non-tumor tissue of patients with cancer (Non-Tumor) and from tumor tissue of patients with cancer (Tumor)

The MSAC results coincides with our understanding of *H. pylori* infection, whereby the tissue is typically neutral⁷⁰ in healthy patients and changes during the progression of inflammation and cancer to become more acidic displaying sialylated and sulfates structures⁷⁸. We believe that the availability of a greater number of MUC5AC samples can be used to ascertain a deeper understanding between *H. pylori* infection and the gastric lining, especially at an oligosaccharide structural level.

3.3 Biological application of MS based glycomic analysis

Characterization of the oligosaccharides is the initial step in understanding the biological relevance. It is important to make the connection between oligosaccharide structure and its biological function. There is little value in describing the glycosylation if this connection cannot be made. We therefore applied our LC-MS platform to analysing *O*-linked oligosaccharides from MUC5AC and *N*-linked oligosaccharides from a secreted fusion glycoprotein with the aim of identifying biologically significant information about their respective glycosylation.

Identification of lacdiNAc: Possible implications in *Helicobacter pylori* infection?

The stomach mucin can act as a functional binding site for certain pathogens such as *H. pylori*. *H. pylori* adheres to the mucus layer via its lectin like adhesins. The best characterised of these adhesins are the Blood group antigen binding Adhesin (BabA), which binds to Le^b and H-Type 1 structures¹⁸⁵ and the sialic acid binding Adhesin (SabA) that binds to sialyl-Le^x and sialyl-Le^a structures⁷⁷. However, other epitopes present on the mucosal surface are involved in inhibiting the binding of *H. pylori*, notably the anti-bacterial GlcNAc α 1-4Gal⁸⁰. In Paper V we investigated the relationship between the *H. pylori* strain J99 and its interactions with the *O*-linked oligosaccharides from MUC5AC of HGM-1 (normal tissue isolated from an individual with gastric cancer) and HGM-2 (healthy individual). J99 was shown to interact differently when cultured on HGM-1 and HGM-2. HGM-1 was found to bind to, and promote, proliferation of J99 in a BabA/Le^b-dependent manner. Interestingly, HGM-2 was negative for both characteristics and did not have any apparent Le^b activity. This was despite the fact that the mucin had been isolated from a secretor individual. Apart from the abundant H Type 2 and H Type 3 structures identified on the HGM-2 which were also present on HGM-1, we found a prominent ion with an [M-H]⁻ m/z 936 on HGM-2 which was in stark contrast to HGM-1 where it was in very low abundance.

We isolated this particular ion for fragmentation and our interpretation of the MSⁿ spectra revealed that it corresponded to a core 2 structure of Gal β 1-3(GlcNAc β 1-6)GalNAc with the C-3 branch terminating with a Fuc residue giving a blood group H Type 3 sequence (Fuc α 1-2 Gal β 1-3GalNAc) and the C-6 branch terminating with the addition of a HexNAc residue (Paper V, Supplementary Figure 2).

This particular fragmentation pattern was interesting since the presence of a terminating diHexNAc sequence is unusual in mucin type *O*-linked oligosaccharides; although it is not unprecedented in *O*-linked glycosylation as structures containing the chitobiose (GlcNAc β 1-4 GlcNAc β 1-)¹⁸⁶ and lacdiNAc (GalNAc β 1-4 GlcNAc β 1-)¹⁸⁷ epitopes have previously been identified on *O*-linked oligosaccharides. Since this particular oligosaccharide was the only distinct feature on HGM-2, we were interested in further characterizing it to identify the terminal HexNAc residue and linkage position.

To achieve this we employed a combination of MS analysis and enzymatic digestion. To determine the terminal HexNAc, we adopted an MS³ approach whereby the fragmentation spectra of known standards containing a terminal GlcNAc β 1-4 and GalNAc β 1-4 were compared to the fragmentation spectra of the dihexNAc sequence identified on the *m/z* 936 structure. In the MS² of the *m/z* 936 structure, we identified a cross ring ^{0,2}A_i fragment with an *m/z* 304 corresponding to the C-4 extension of the core 2 GlcNAc which was also present in the MS² spectra of both of the standards. Since this particular fragment retained both the full terminal HexNAc residue as well as linkage position we isolated it from the HGM sample and the two standards for MS³ fragmentation. We were able to correlate the MS³ fragments and their respective intensities from both standards with the MS³ fragments intensity of the diHexNAc structure from the gastric sample. The standard with a terminal β 1-4GlcNAc had a R² value of 0.49 compared to an R² value of 0.95 for the standard with a terminal β 1-4GalNAc. This confirmed that the terminal HexNAc residue was a C-4 linked GalNAc (Figure 12).

With having identified the terminal HexNAc to be a C-4 linked GalNAc, we investigated whether it was α or β linked by digestion using exo *N*-acetylhexosaminidases. There were limitations in the amount of HGM available for digestion, however we were able to identify the same [M-H]⁻ *m/z* 936 dihexNAc structure on porcine gastric mucin (PGM) (Paper V, supplementary Figure 1) and therefore we chose to digest PGM as well as HGM with the exo *N*-acetylhexosaminidases. We treated the *m/z* 936 structure with α -*N*-acetylgalactosaminidase that removes α linked galactose, a β -hexosaminidases that removes β -2,4 and 6 linked GlcNAc and a β -hexosaminidases that removes both β linked GlcNAc and GalNAc residues and compared the intensity of the *m/z* 936 structure before and after digestion to the intensity of an internal standard that was not affected by the treatment. The results of the digestion revealed that the terminal GalNAc was β 1-4 linked (Paper V, Figure 4) as the *m/z* 936 structure was only digested by the β -hexosaminidases that removes both β linked GlcNAc and GalNAc residues. With these results we were able to conclude that the terminal structure was a lacdiNAc (GalNAc β 1-4GlcNAc) epitope.

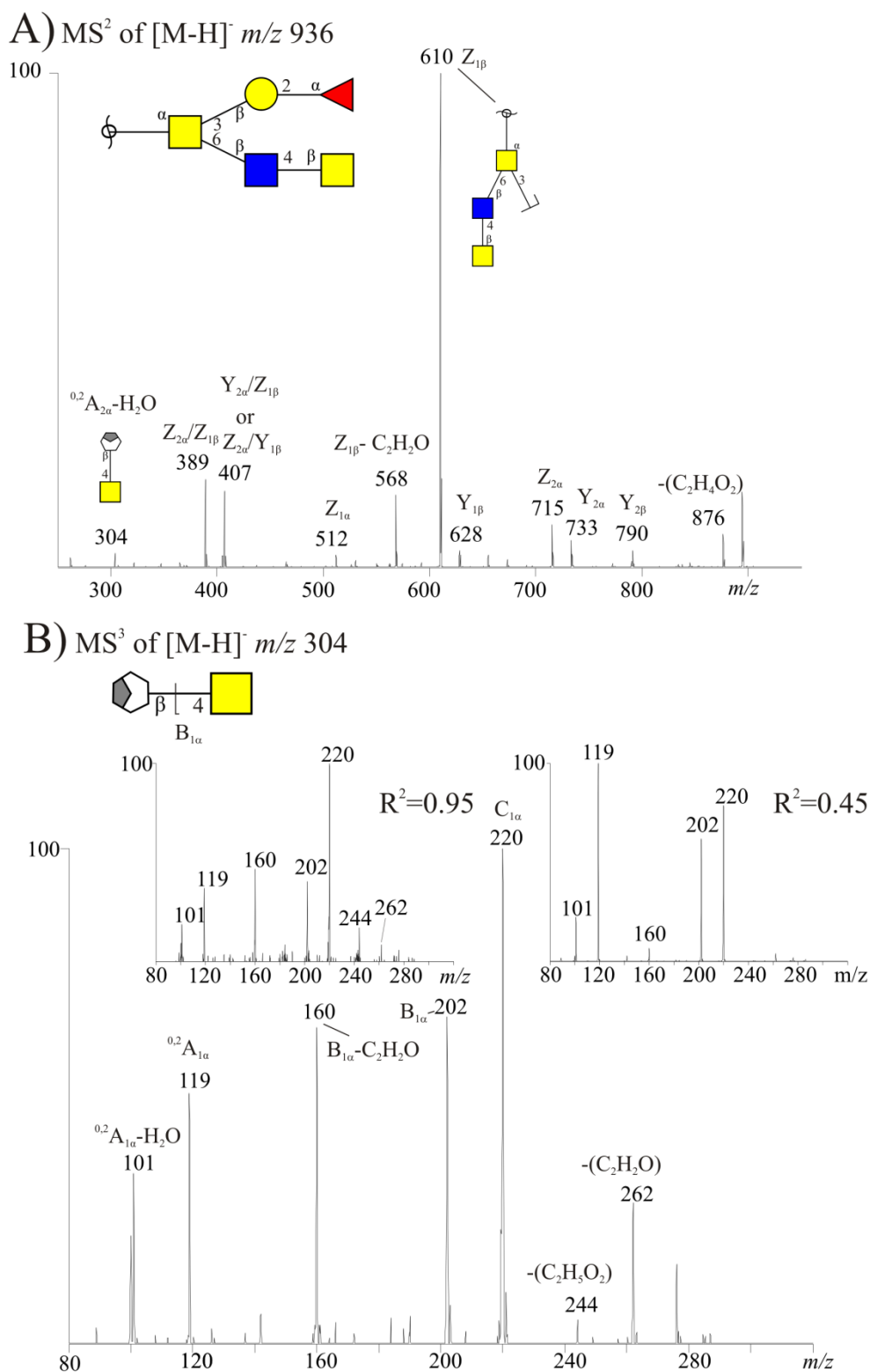


Figure 12: Identification of the lacdiNAc sequence by MSn. The MS² fragmentation of the component Fucal-2Galβ1-3(GalNAcβ1-4GlcNAcβ1-6)GalNAcol to identify the C-6 branch of the GalNAcol ($Z_{1\beta}$ fragment ion of m/z 610) and the linkage position of the GalNAcβ1-4GlcNAcβ1- moiety ($^{0,2}A_{1\alpha}-H_2O$ C-4-specific fragment ion of m/z 304) (A). (B) The MS³ fragmentation of the $^{0,2}A_{1\alpha}-H_2O$ ion is seen containing the GalNAcβ1-4 moiety plus part of the cleaved GlcNAc. Inserts show the MS³ fragmentation of the $^{0,2}A_{1\alpha}-H_2O$ fragment ion of m/z 304 isolated from GalNAcβ1-4Gal (left) and GlcNAcβ1-4GlcNAcβ1-4GlcNAc, showing with the r^2 values that the linkage of configuration of the sample correspond to the GalNAc1-4 linkage of the standard.

Does lacdiNac inhibit the synthesis of Lewis epitopes?

With the identification of the terminating lacdiNac epitope as a major structure present on HGM-2, we believe that it is possible that the lacdiNac epitope is involved in the binding of the *H. pylori* strain J99 to MUC5AC. The α 1-2 fucosyltransferase (the secretor gene product) and GalNAc β 1-4 transferase are clearly very active during *O*-linked synthesis. Both of these transferases share the common feature of preventing poly-*N*-acetylglucosamine extension of the C-6 arm. This is the same branching arm that would normally express other terminal epitopes such as Le^b and sLe^x epitopes.

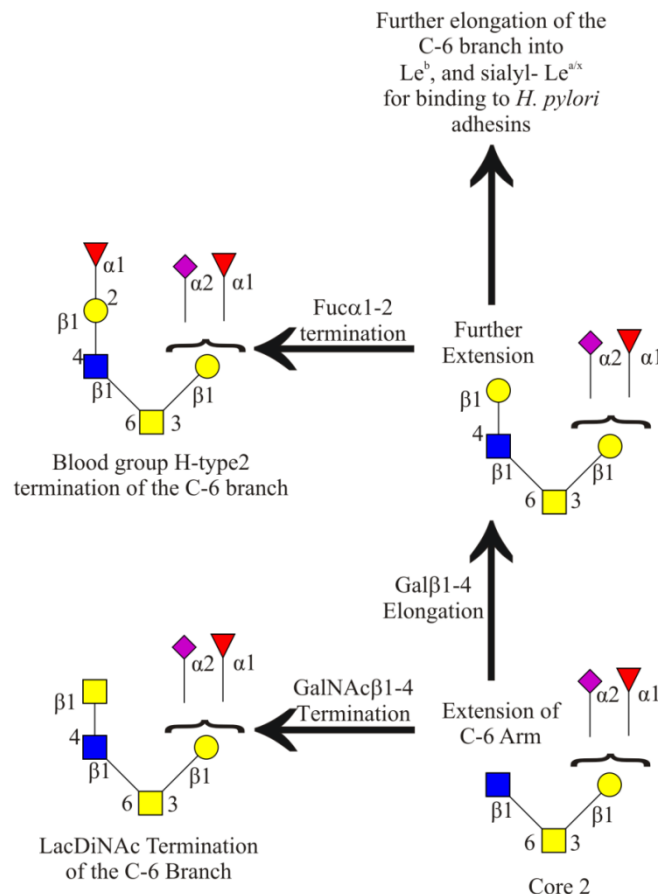


Figure 13: Prevention of the biosynthesis of complex core 2 mucin oligosaccharides by the secretor gene (α 1-2 fucosyltransferase) and the β 1-4 *N*-acetylgalactosamine transferase by terminating the extension of the C-6 branch.

It is possible therefore, when these terminating glycosyltransferases are highly active during oligosaccharide synthesis, they limit the availability of core 2 structures to act as acceptors for the other blood group and Lewis associated glycosyltransferases, including Le^b and sLe^x. This could limit the synthesis of Le^b and sLe^x structures, which are necessary for *H. pylori* adhesion via BabA and SabA respectively. This could explain why there is limited diversity in the type of oligosaccharide present on HGM-2 with fewer extended structures, whereas on HGM-1 where the GalNAc β 1-4 transferase activity is much lower we see greater diversity in the types of structures present as a greater number of core 2 structures undergo further extension of the C-6 branch (Figure 13).

To investigate this hypothesis further, we also analysed the additional MUC5AC samples isolated from healthy individuals and tumor and non-tumor sample from patients with gastric cancer (used to generate *O*-linked MSAC data, Figure 11) with the *Wisteria floribunda* lectin (lacdiNAc) and the anti-Le^b monoclonal antibody (as per the method described in Paper V). The results show an interesting correlation between the presence of lacdiNAc in tumor and non-tumor cancer samples with a general decrease in the amount of lacdiNAc in the tumor samples compared to the non-tumor samples (Figure 14). We believe this decrease is possibly a result of increasing amounts of core 2 oligosaccharides being terminated by sialyl Le^{a/x} epitopes instead of lacdiNAc (and possibly also Le^b epitopes). Interestingly, it was noted that the healthy sample from the secretor individual was negative for the presence of Le^b structures, while showing a strong response for the presence of lacdiNAc.

The MSAC data for the MUC5AC samples in Figure 11 shows increasing sialic acid in tumor samples, which could be a result of increasing amounts of sLe^x displaying structures. This would fit with our model of *H. pylori* infection (Figure 3 of the introduction) where the gastric mucin becomes more acidic (with increasing amounts of sialylated and sulfated oligosaccharides) with the onset of inflammation and cancer. This will require further investigation; however these initial results support our hypothesis that by acting as a terminating epitope, lacdiNAc may limit the addition of other terminating epitopes such as Le^b and sLe^x and by doing so indirectly contributes to the inhibition of the binding of *H. pylori* to the gastric mucus layer.

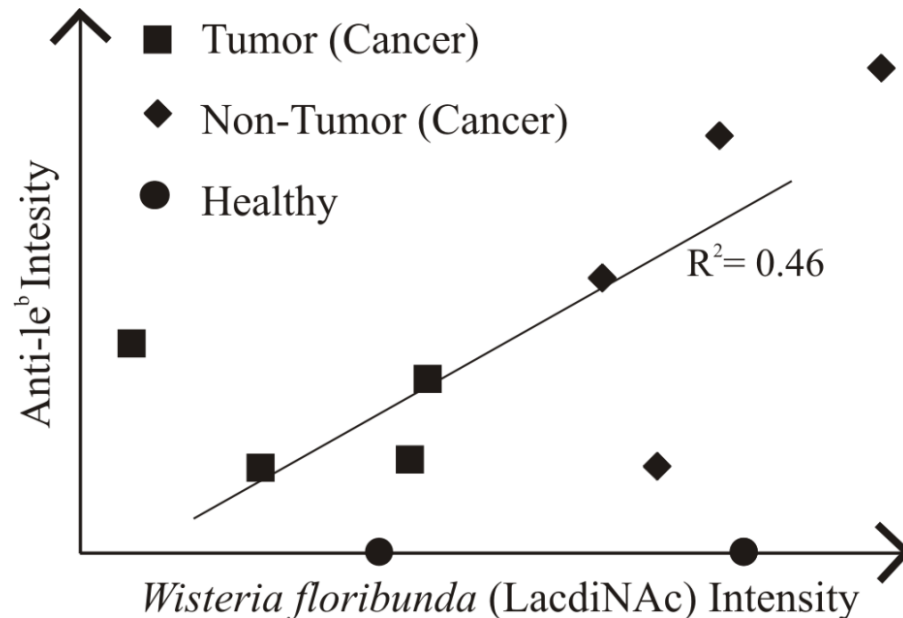


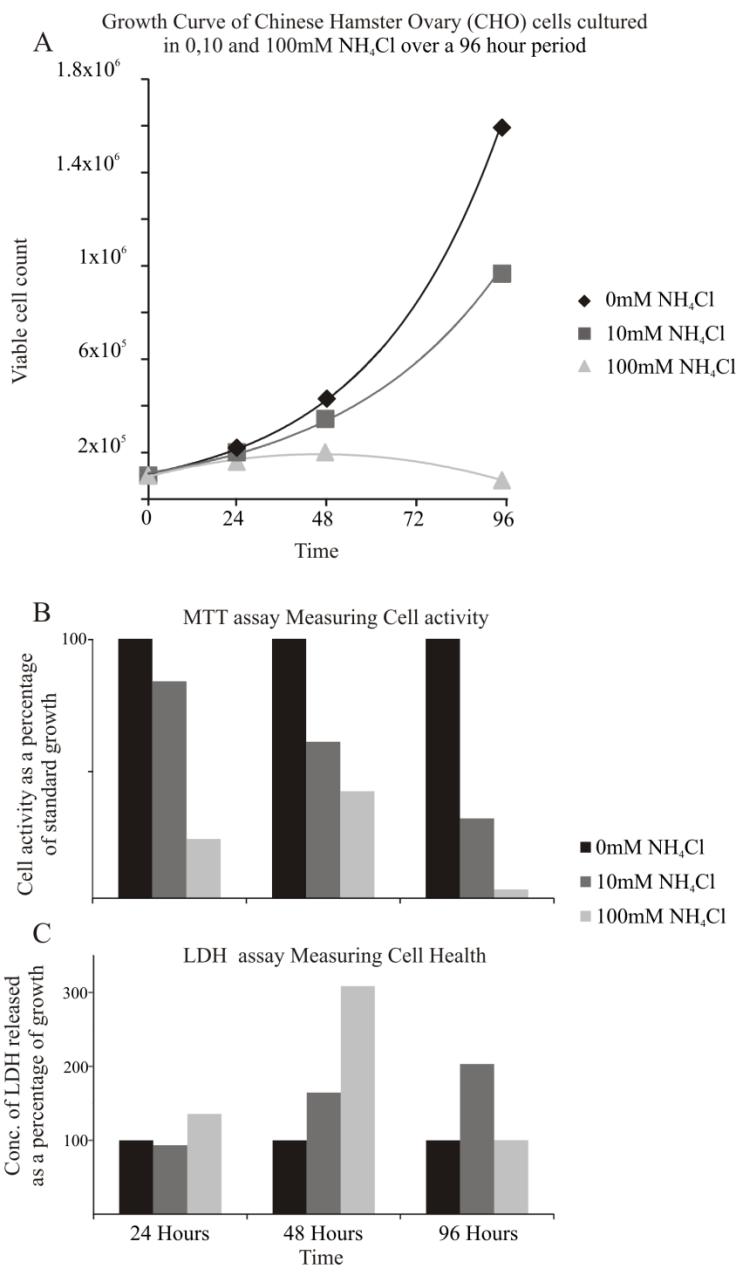
Figure 14: Human gastric MUC5AC which was isolated from healthy individuals and from the non-tumor tissue and tumor tissue of individuals with gastric cancer was blotted onto a PVDF membrane and probed with biotinylated lectin from *Wisteria floribunda* and a monoclonal antibody against Le^b.

The importance of detecting aberrant glycosylation

In paper IV we used LC-MS to characterize the glycosylation of membrane associated oligosaccharides with the aim of comparing the samples to highlight differences in their glycosylation. In particular, using LC-MS, we could show that the oligosaccharide composition of the two membranes samples differed significantly with one membrane containing high amounts of sialylated structures while the other samples contained mainly unsialylated structures. Detecting differences such as changes in the level of sialylation is important as it could indicate aberrant glycosylation which can have profound biological consequences; increasing levels of sialylation, for example, can serve as a biomarker for certain cancers^{188,189}. Furthermore sialylation acts as a self-determining agent for the body's immune system¹⁹⁰. This function, in particular, is of increasing interest in the biotechnology industry where the level of sialylation can impact the pharmacokinetic properties of glycosylated biopharmaceuticals³⁴. Undersialylation can be a particular problem for efficacy of a biopharmaceutical product as asialyloglycoproteins are selectively removed in the liver via the asialoglycoprotein receptor, thus reducing their *in-vivo* half-life¹⁴. Therefore a homologous glycosylation profile is desirable when manufacturing glycosylated biopharmaceutical products. However, it is unfortunately rarely achieved throughout the duration of a batch fermentation cycle due to the introduction of toxins in the culture medium.

Ammonium ions (NH_4^+), a by-product of glutamine metabolism, in particular have been identified as a causative agent in improper glycosylation, as in the ammonia (NH_3) form it can cross the trans-golgi membrane and alter inter-golgi pH, thus causing redistribution of the glycosyltransferases⁴¹ and has been correlated with a decrease in terminal sialylation⁴². Concentrations of around 10mM of NH_4^+ have been reported towards the end of a typical batch fermentation cycle¹⁹¹.

We could show that these higher concentrations significantly impact the cell health, reducing cell activity and lowering the cell culture growth rate of a CHO cell line that secretes a fusion glycoprotein. The addition of 10mM of NH_4Cl into the culture media reduced the growth rate by over 40% and cell activity by over 60% after a 96 hour incubation period. Furthermore the increasing secretion of lactate dehydrogenase into the media suggests that the presence of NH_4Cl impacts on the cells membrane integrity (Figure 15). The fusion glycoprotein secreted from the CHO cells is composed of the Fc region of the immunoglobulin IgG1 and the extracellular matrix of CTLA-4¹⁹². Identifying aberrant glycosylation of this protein is important since the *N*-linked oligosaccharides are essential for the *in-vivo* stability of the active CTLA-4 portion of the fusion protein¹⁹³. It is therefore important to monitor the glycosylation profile of the secreted fusion protein during production to ensure that it is correctly glycosylation.



32

Figure 15: Growth Characteristics of Chinese Hamster Ovary CHO cultured in the presence and absence of ammonium chloride. A) The Growth curve of CHO cells cultured over a 96 hour period in serum free media substituted with 8mM glutamine at 37°C in 5% CO₂ with the addition of 0mM, 10mM and 100mM of NH₄Cl. B) Bar Chart representing the Cellular activity as measured by the MTT assay. The results were normalised so that cells cultured under normal conditions (i.e. 0mM NH₄Cl) represented 100% of expected cellular activity, thus decreased activity is represented as percentage below the 100% threshold. C) A bar chart representing cellular health as indicated by the presence of lactate dehydrogenase (LDH) in the culture media leaked from damaged cells. The results were normalised so that cells cultured under normal condition represented 100% of expected LDH leakage, thus increased LDH leakage is represented as a percentage above the 100% threshold.

The concentration of NH_4^+ cations present in the culture media increases throughout the fermentation cycle, with relatively low concentrations present during the initial stages and the concentration generally increasing towards end of the fermentation cycle. Analysis of *N*-linked oligosaccharides of the fusion glycoprotein harvested during the initial and the latter stages of a fermentation cycle show there is no significance in the relative amount of biantennary sialylated structures with one ($[\text{M-H}]^-$ m/z 1038) versus two ($[\text{M-H}]^{2-}$ m/z 1183) sialic acid residues present on the glycoproteins harvested in the initial and latter stages of the fermentation cycle (see Figure 16A and 16B). However, the data showed the presence of the high-mannose Man 5 ($[\text{M-H}]^-$ m/z 1233) structure only on the glycoproteins harvested at the latter stage (see Figure 16C and 16D).

To further investigate the prominence of the high-mannose structure, the protein content and the mannose content of each fraction was determined by probing with an antibody specific to the protein and with the lectin Concanavalin A that binds to mannose residues. The ratio of mannose residues to protein concentration was far higher in the latter fraction (1.68) compared to the early fraction (0.77) (see inserts Figure 16C & 16D) which correlated to a high abundance of mannose on the glycoprotein. The presence of the high-mannose structure on the protein harvested during the latter stages of the fermentation cycle is probably a result of incomplete processing of the *N*-linked oligosaccharides in the Golgi apparatus after initial processing in the endoplasmic reticulum. The data suggests incomplete processing of nascent oligosaccharide and indicates that the Golgi apparatus is affected by the treatment. Determining the point at which this incomplete processing occurs is important for quality control process as it can determine the optimal time to harvest the glycoprotein and LC-MS provides a fast and efficient method of detecting the aberrant glycosylation.

We have attributed the presence of the high-mannose structure to the increasing toxicity associated with NH_4Cl , however we have considered the possibility that there could have been other factors within the bioreactor that contribute to the alteration in glycosylation of the expressed fusion glycoprotein such as shear stress or glucose starvation. It has been shown that increasing shear stress will result in a decrease in glycosylation site occupancy, however the shear stress required was considered to be very high and not what would be expected in a typical bioreactor¹⁹⁴. Extreme glucose starvation has been implicated as affecting site occupancy and glycosylation, however even though glucose depletion leads to a reduction in UDP-GlcNAc concentrations, relatively large fluctuations in the concentration resulted in minor changes in glycosylation¹⁹⁵.

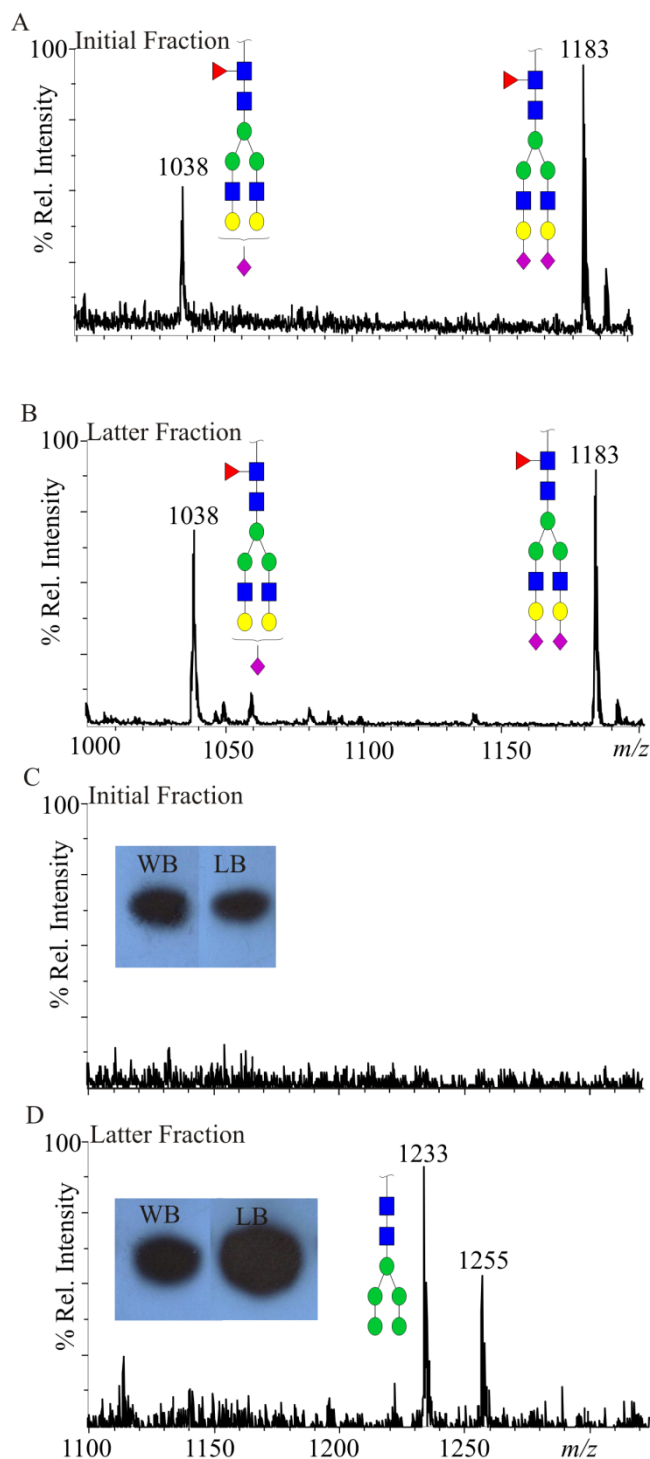


Figure 16: Identification of complex and high mannose type *N*- linked oligosaccharides isolated from a glycoprotein secreted from CHO cells. Two complex structures, $\text{Man}_3\text{GlcNAc}_4\text{Fuc}_1\text{Gal}_2\text{NeuAc}_1$ and $\text{Man}_3\text{GlcNAc}_4\text{Fuc}_1\text{Gal}_2\text{NeuAc}_2$, were present on the glycoprotein harvested at the early and latter stages of a fermentation cycle (16A & 16B). The high-mannose structure was not detected on the glycoprotein harvested at the early stage but present on glycoprotein harvested in the latter stage of the fermentation cycle (16C & 16D). Insert in Figure 16C and 16D shows a western blot with a mouse anti-IgG1 antibody to determine the concentration of the protein and a lectin blot with Concanavalin A to determine the concentration of the mannose residues.

4 Conclusion

The overall goal of this thesis was to adopt our LC-MS platform for glycomic analysis of glycoproteins. In particular we were interested in characterizing complex glycomic samples such as membrane associated *N*-linked oligosaccharides and the mucin *O*-linked oligosaccharides with the aim of identifying particular glycosylation traits or glyco-epitopes that could have biological significance.

The application of mass spectrometry to the analysis of biomolecules is well established and MS is at the forefront of our glycomic platform. LC-MS, in particular, is well suited to this purpose as there is generally a mix of different glycoforms present on a glycoprotein and the ability to separate these out prior to MS analysis improves the data acquisition. For structural assignment of a particular oligosaccharide, especially where information about linkage and/or branching is important, fragmentation by CID and subsequent manual interpretation of the MSⁿ spectra is favoured. CID is a well-established fragmentation pathway used in glycomics as it produces predictable fragmentation patterns and provides information rich MSⁿ spectra.

However, despite these excellent characteristics, CID is not without its complications in structural assignment of oligosaccharides. The rearrangement of fragment ions during CID can complicate the interpretation of the MSⁿ and can lead to mis-assignment of an oligosaccharide structure. The migration of fucose during CID fragmentation in positive ion mode has been known for over a decade. However, in this thesis we have shown that the migration of sulfate groups can also occur when sulfated oligosaccharides are fragmented by CID in negative ion mode. Our initial observation was that sulfate migration occurred in the [M-H]⁻ ions of sulfated-sialylated structures and that it was absent in [M+Na-H]⁻ and [M-2H]⁻² ions which indicated that the migration was promoted by the interaction between the ions and a mobile proton within the ion trap. Further analysis of sulfated oligosaccharide revealed that a secondary condition of increased conformational freedom of the donating and acceptor monosaccharide residue increased the degree of sulfate migration. Although we have shown that fragmentation by HCD limits the migration of sulfated residues to where it cannot be detected, CID is a more universally used fragmentation pathway and HCD may not be available in every laboratory. Therefore, it may be that the true value of the results generated in this thesis is to make the glycobiologist aware that migration of sulfate can occur and that this should be considered when interpreting MSⁿ spectra of sulfated oligosaccharides in negative ion mode.

One aspect we found particularly challenging was how best to display the MS data in such a manner that could be easily interpreted by both the experienced glycobiologist, who is well acquainted with MS derived data, as well as other scientists whose expertise lies outside the field of glycomics.

While developing our membrane preparation protocol for the glycomic analysis of membrane associated *N*-linked oligosaccharides we were interested in exploring how to optimize our methods of displaying our data. We found a table of identified structures was an unsatisfactory approach as it provided limited information and was particularly hard for the non-glycobiologist to digest. We have also found that the generation of semi-quantitative data alone was unsatisfactory as it did not provide an overview of the glycosylation profile nor make the connection between oligosaccharide structure and biological function. In our study, we have found that a process of sorting the oligosaccharides by compositional analysis (such as MSAC) to provide a global overview of the membrane glycosylation or by biological characteristics (such as common core extensions or terminal epitopes) provided a better overview of the biological function of the oligosaccharides. Our efforts to display the MS data was by no means exhaustive however it does highlight the importance of developing strategies for effective dissemination of glycomic data

The importance of opening up glycomic data to a wider audience should lead to greater interdisciplinary studies which will hopefully help bridge the gap between identifying particular glyco-epitopes and understanding their biological importance. Our analysis of *O*-linked oligosaccharides derived from MUC5AC of healthy and tumor-associated tissue has identified the presence of the lacdiNAc epitope on gastric mucin oligosaccharides. We believe this epitope plays a role in inhibiting the binding and proliferation of *H. pylori* to the gastric mucus layer. However, whereas the GlcNAc α 1-4Gal- epitope is directly involved in inhibiting binding through its anti-microbial properties, it is more probable that lacdiNAc's involvement is indirect in so far as it limits the synthesis of oligosaccharide displaying Le^b and sLe^x epitopes which are required for the adhesion of *H. pylori* via its adhesins BabA and SabA respectively. The specific role that the lacdiNAc epitope plays in *H. pylori* infection will need to be investigated further and provide a greater insight into the complex relationship between *H. pylori* infection and the gastric mucins.

5 Future Perspectives

Our strategy of characterizing the oligosaccharides relies upon the manual interpretation of the MSⁿ spectra for structural elucidation. While the acquisition of information-rich MSⁿ spectra is generally important for data analysis, for manual interpretation of MSⁿ spectra it is essential. The availability of different fragmentation pathways such as HCD, ETD (electron transfer dissociation), EDD (electron detachment dissociation) amongst others, which are complimentary to CID, can help to improve the acquisition of glycomic MS data. We have identified HCD as a particularly useful fragmentation pathway for the analysis of sulfated oligosaccharides as it limits the migration of the sulfate residue while avoiding neutral loss of the sulfate groups. Furthermore ‘stepped HCD’, a process by which the parent ion is fragmented with two different collision energies and the resulting daughter ions from both fragmentation events are detected in the same scan, could be a useful tool at providing fragment ions from both cross ring and glycosidic bond cleavage events.

While better acquisition of MSⁿ spectra should improve the process of interpreting MS data, manual interpretation remains an impediment to rapid processing of MS data. Increasing the availability of MS based computational software suitable for glycomic analysis would clearly be beneficial for the field of glycomics. The incorporation of MSⁿ spectra from different fragmentation pathways into existing glycomic databases could be the first step in this process. Ideally glycomic based MS analysis would move towards automating the processing of MS data similar to what has happened in MS based proteomic analysis.

Improving the characterization of oligosaccharides should aid investigations into the biological role of glyco-epitopes. MS was key to our identification of the lacdiNAc epitope present on gastric mucin. Our initial analysis shows that this particular epitope could be important in inhibiting the binding of *H. pylori*. Since *H. pylori* is considered the primary cause of gastric cancer, there is clearly an important clinical value in understanding the complex relationship between it and the gastric mucus layer. Our initial study was limited to samples from two individuals, however with greater availability of samples we hope to improve on the data generated to date, which should provide a more robust investigation of the gastric mucin *O*-linked oligosaccharides. Furthermore by incorporating the MS results into freely available databases, such as Unicarb-DB, it is hopeful that other research groups investigating *H. pylori* infection can utilise the MS data to improve both our understanding of *H. pylori* infection and our efforts to cure it.

6 Supplementary Materials and Methods

All materials were obtained from Sigma Aldrich (St Louis, MO) unless otherwise stated. The solutions were prepared using 18-m Ω water, which was produced using the MilliQ water purification system (Millipore, Billerica, MA)

Release of O-linked oligosaccharides for permethylation and MS analysis

The glyco-variant of MUC5AC isolated as HGM was blotted onto Immobillin P PVDF membranes (Millipore) and stained with DB71 and destained in 10% acetic acid in 40% ethanol. The *O*-linked oligosaccharides were released in 0.5 M NaBH₄ in 50 mM NaOH. The reaction was suspended with acetic acid and the samples were desalted with 60 μ L of AG 50W-X8 cation exchange beads (Bio-Rad, Hercules, CA) packed in C18 zip tips (Millipore). Borate complexes were removed by repeated addition/evaporation with 1% acetic acid in methanol (100 μ L for each addition).

Permethylation of release *O*-linked oligosaccharides

O-linked oligosaccharides from human gastric mucin derived MUC5AC were per-*O*-methylated as described by Kerek et al¹⁵⁰. Briefly, the desalted oligosaccharides were dried and dissolved in 500 μ L of dimethyl sulfoxide, followed by the addition of powdered sodium hydroxide and mixed by vortexing. 100 mL of iodomethane was added. The sample tubes were capped and wrapped in aluminium foil and incubated on a shaker at room temperature for 1 hour. After incubation, the sample was placed on ice and the reaction was stopped by the addition of 1 mL of water and 1 mL of dichloromethane and mixed by vortexing. The top organic phase (containing the per-*O*-methylated oligosaccharides) was removed and transferred to a clean tube. Dichloromethane addition was repeated twice and each time the top organic phase was extracted and pooled. The new tube with the sample was washed with HPLC grade water until the solution was pH 7 to ensure the sodium hydroxide had been removed. During washing, the water layer was removed between each wash and discarded. The dichloromethane was evaporated by speed vacuum.

Mass Spectrometric analysis of Per-*O*-methylated *O*-linked oligosaccharides

Per-*O*-methylated oligosaccharides were analysed by direct infusion into an Orbitrap mass spectrometer. The dried oligosaccharides were dissolved in methanol and transferred to a Type F thin wall nanoflow probe tip (Waters, Milford, MA) and ionised by ESI in positive ion mode with a spray voltage of 1.6 kV and detected by an Orbitrap XL mass spectrometer.

Fragmentation of the selected ions was by collision induced dissociation with a collision energy of 35%. MSⁿ spectra was accumulated over a 1 minute time frame with a mass range of *m/z* 200-2000.

Analysis of N-linked oligosaccharide released from a glycoprotein secreted by a Chinese Hamster Ovary Cell

Culture of CHO cells

CHO Cells (CRL-10762, ATTC) were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 0.1% FBS after weaning the cells from an initial concentration of 10%FBS. For maintenance the cells were cultured in T flasks with 10mL of media per 25 cm² surface area and incubated at 37°C, buffered in 5% CO₂. The flasks were passaged every 3 days.

Determination of the culture growth curve

The growth curve of CHO cells cultured with and without the addition of ammonium chloride was determined. T-25 flasks (corning) were seeded with 1 x 10⁵ cells in 10mL of media with the addition of 0, 10 or 100 mM of ammonium chloride. The flasks were incubated for a total of 96 hours at 37°C in 5% CO₂. After each 24 hour period of incubation, the viable cells count for each concentration of ammonium chloride was determined in triplicate (therefore, in total 12 flasks are required for four time points per test condition). Briefly, The media was removed from the flasks and the flasks were washed with 10mL of phosphate buffered saline (PBS). The viable cells were removed from the flask by treatment with trypsin-EDTA and a viable cell count was determined by counting in a haemocytometer after trypan blue staining. The media from the remaining flasks were removed and discarded. The flasks were washed with 10mL of PBS and 10 mL of fresh media with the appropriate concentrations of ammonium chloride was added to the flasks and the flasks re-incubated.

Determination of cellular activity by MTT assay

The cellular activity of CHO cells cultured with and without ammonium chloride was determined by an MTT assay. CHO cells were cultured in a 96 well plate with a seeding density of 2x10⁴ cells/cm² and 0, 10 or 100 mM of ammonium chloride was added to the media. The cells were incubated at 37°C in 5% CO₂ for a total of 96 hour period. After each 24 hour period of incubation the cellular activity was determined by MTT assay in triplicate. 20µL of MTT reagent was added to the appropriate well and also a well containing only culture media as a negative control. The well-plate was incubated at 37°C for 2-4 hours on an orbital shaker. After incubation, the absorbance was measured at 450 nm and subtracted from the reference wavelength measured at 690 nm by a spectrophotometer.

Determination of cell health by LDH assay

The cell health of CHO cells cultured with and without ammonium chloride was determined by a LDH assay which measures the amount of lactate dehydrogenase that leaks from the cell due to damage from the cellular membrane. CHO cells were cultured in a 96 well plate with a seeding density of 2×10^4 cells/cm² and 0, 10 or 100 mM of ammonium chloride was added to the media. The cells were incubated at 37°C in 5% CO₂ for a total of 96 hour period. After each 24 hour period of incubation the cell health was measured by LDH assay in triplicate. The cell media was removed to 0.5mL centrifuge tubes and centrifuged for 4 minutes at 250g to remove cell debris. 100µL of supernatant was transferred to a fresh tubes and 200µL of LDH reagent was added to the supernatant and a tube with fresh media which was the negative control. The tubes were vortexed, covered in aluminium foil and incubated at 37°C for 30 minutes. After incubation the reaction was stopped by the addition of 20µL of 1M HCL. The absorbance was measured at 490 nm and subtracted from the reference wavelength measured at 690 nm by a spectrophotometer.

Release of *N*-linked oligosaccharides for LC-MS analysis

The solubilised membrane extracts were blotted onto Immobilon-P PVDF membranes (Millipore, Billerica, MA) and stained with DB71 and de-stained in 10% acetic acid in 40% ethanol. The membranes were blocked with 1% polyvinylpyrrolidone (PVP) in methanol. The *N*-linked oligosaccharides were enzymatically released from the protein by PNGase F. The enzyme (5µL, 5 units) was added to each blot and incubated for 10 minutes at 37°C. After 10 minutes incubation a further 5 µL of PNGase F and 10 µl of water was added and the blots were incubated overnight at 37°C. The released oligosaccharides were reduced by the addition of 0.5 M NaBH₄ in 50 mM NaOH and incubated at 60°C overnight. The reaction was suspended with acetic acid and the samples were desalted with 60 µl of AG 50W-X8 cation exchange beads (Bio-Rad, Hercules, CA) packed in C18 zip tips (Millipore, Billerica, MA). Borate complexes were removed by repeated addition/evaporation with 1% acetic acid in methanol (100 µL for each addition). The released oligosaccharides were dissolved in water prior analysis by LC-MS.

LC-MS of released *N*-linked oligosaccharides

Released membrane associated oligosaccharides were analysed by LC-MS using a 10 cm × 250 µm I.D. column containing 5-µm porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA) prepared in-house¹⁹⁶. Oligosaccharides were eluted using a linear 8mM aqueous ammonium bicarbonate/ acetonitrile gradient from 0-40% in 40 minutes at a flow rate of 10 µL/minute. The eluted oligosaccharides were analysed using an ESI-IT MS (LTQ, Thermo Electron Corp., San Jose, CA) operating in negative ion mode with a spray voltage of -3.5 kV. Oligosaccharides were detected as [M-H]⁻, [M-2H]⁻² and [M-3H]⁻³ over a scan range of *m/z* 380-2000. Individual oligosaccharides were isolated for fragmentation by collision induced dissociation (CID) with the collision energy of 35% and a dynamic integration time.

Lectin and antibody probing of HGM-1, HGM-2 and PGM.

The enriched fusion protein samples harvested at the start and end of a fermentation batch cycle were provided by the Centre for Bioanalytical Science (CBAS), National University of Ireland, Galway. 10µg of the enriched glycoprotein samples were transferred to a PVDF membrane. Porcine Gastric Mucin (PGM) was also transferred and used as a negative control. The membranes were blocked with 1% bovine albumin in Tris Buffered Saline with 0.05% Tween 20 (TBST). The Blots were probed with a biotinylated lectin from Concanavalin A (Vector Laboratories, Burlingame, CA) diluted to 1:2,000 in TBST and mouse anti-IgG1 mono-clonal anti-bodies (Sigma Aldrich, St Louis, MO) diluted to 1:1,000 in TBST for 3 hours incubation at room temperature. After probing, the blot were washed 5 times with TBST for five minutes. The blots were incubated in horseradish peroxidase conjugated to streptavidin diluted to 1:8000 in blocking solution for blots probed with lectin from Con. A and with horseradish peroxidase conjugated to goat anti-mouse IgG (P0161, Dako, Glostrup, Denmark) diluted to 1:20,000 in blocking solution for blots probed with the IgG antibody for 60 minutes at room temperature. The blots were further washed for 5 times with TBST and once with TBS. The blots were developed for 1 minute using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Waltham, MA) and the blots. signal intensity was measured using ImageJ software (National Institute of Health, Bethesda, MA, USA).

7 Acknowledgements

There are a lot of people who I would like to thank for their support throughout my PhD. There are far too many to mention in one or two pages and they far too important to forget. This thesis is dedicated to everyone who made me smile throughout the years of my PhD studies. However there are a few people who deserve specific mention.

Niclas Karlsson for introducing me to the world of mass spectrometry. I had actually never seen a mass spec. instrument before starting my PhD, however through your tutorship I feel I can describe myself as a mass spectrometrist. You have taught me to appreciate the synergy between oligosaccharide structure and biological function. I may not fully believe that ‘if a problem can’t be solved by mass spectrometry, it’s not worth solving’ however I can at least appreciate the ethos. Apart from mentoring me, you have also provided me with many laughs and fun times.

Everyone I have worked with either as a member of CBAS at NUIG or in the Karlsson group at Medkem in GU for the good times. In no particular order or importance I would like to thank **Brendan** for helping me out when I didn’t even know what end of the MS the sample went into. **Liaqat “The Qatt”** and **Jin** for their help with preparing my manuscript and just generally always been willing to help when asked. **Sam** for our various sunshine breaks and gym sessions (but not for stealing my car!!), **Catherine, Roisin, Lucia** and **Siobhán** for the fun times and morning coffee at Moffets. **Prof. Angela Savage** for offering me the PhD with CBAS.

The members of **Niclas’s, Sara’s, Susann’s** and **Gunnar’s** group at Medkem in GU for all the good conversations over fika and making me feel welcome. I would like to mention a few people who provided some great laughs over beers on Friday, again in no particular order. **Sjoerd** and **Daniel** (also thank you for your help with preparing my manuscript) **André, Jenny, Emma, Jayesh, Taher, Karin, Jessica, Tina, Robert, Ana** and **Christian**

Prof. Paul Murphey, Karen Kelly and the rest of the staff and students in the chemistry department at NUIG for their support.

Rob and everyone in the Woods group, either at the Chemistry department in Galway or the CCRC at UGA. **Nina, Elisa** and **Lori** for the fun times at the bar in NUIG. **Mark** and everyone at CCRC for hosting me and making me very welcome. I would like to reserve a special thank you to **Matt** and **Tzafra** for taking me into their home and making me feel like family. I feel blessed to have met you both.

My colleagues at Core Facility in GU. **Carina, Sara, Kanita, Petra, Jörgen** and **Elisabet** for their support while writing my thesis and for keeping me alert with plenty of coffee.

My collaborators **Sara, Emma, Tina, Weston** and **Pauline**

May Ling for encouraging me to do a PhD. If only I knew back then what I know now.....

All my friends for their support and encouragement. Whether they are in Salisbury, Dublin, Galway, GA, Gothenburg or anywhere else in the world. If you can judge a man by the friends he keeps, then I think I can be judged favourably.

My brother **Ciarán** and my sisters **Niamh** and **Clodhna** as well as my nephews **Seán, Oisín** and **Cormac** for just being there. My **Mum** and **Dad** for everything you do for me. I owe you so much and you have my eternal gratitude.

Valeria for your constant support and positive energy and for always reminding me that “It’s going to be good”. Ti amo tanto.

8 REFERENCES

1. Lamond, A.I., *Molecular biology of the cell, 4th edition*. Nature, 2002. **417**(6887): p. 383-383.
2. Varki, A., *Essentials of glycobiology*. Essentials of glycobiology, ed. A. Varki, et al. 1999: Cold Spring Harbor Laboratory Press {a}, 10 Skyline Drive, Plainview, New York 11803, USA. xvii 653p.
3. Taylor, M.E. and K. Drickamer, *Introduction to glycobiology*. Introduction to glycobiology. 2003: Oxford University Press, Great Clarendon Street, Oxford, Oxon, OX2 6DP, UK. i.
4. Elbein, A.D., *The role of N-linked oligosaccharides in glycoprotein function*. Trends in biotechnology, 1991. **9**(1): p. 346-352.
5. Kiani, C., et al., *Structure and function of aggrecan*. Cell Res, 2002. **12**(1): p. 19-32.
6. Loomes, K.M., et al., *Functional protective role for mucin glycosylated repetitive domains*. European Journal of Biochemistry, 1999. **266**(1): p. 105-111.
7. Mudad, R. and M.J. Telen, *Biologic functions of blood group antigens*. Current Opinion in Hematology, 1996. **3**(6): p. 473-479.
8. Lisowska, E., *The role of glycosylation in protein antigenic properties*. Cellular and Molecular Life Sciences, 2002. **59**(3): p. 445-455.
9. Endo, T., *Aberrant glycosylation of alpha-dystroglycan and congenital muscular dystrophies*. Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology / edited by the Gaetano Conte Academy for the study of striated muscle diseases, 2005. **24**(2): p. 64-9.
10. Collins, J. and C.G. Bonnemann, *Congenital Muscular Dystrophies: Toward Molecular Therapeutic Interventions*. Current Neurology and Neuroscience Reports, 2010. **10**(2): p. 83-91.
11. Jaeken, J. and H. Carchon, *Congenital disorders of glycosylation: a booming chapter of pediatrics*. Current Opinion in Pediatrics, 2004. **16**(4): p. 434-439.
12. Schachter, H., *Congenital disorders involving defective N-glycosylation of proteins*. Cellular and Molecular Life Sciences, 2001. **58**(8): p. 1085-1104.
13. Wang, F.-L., et al., *High expression of alpha 2, 3-linked sialic acid residues is associated with the metastatic potential of human gastric cancer*. Cancer Detection and Prevention, 2009. **32**(5-6): p. 437-443.
14. Hossler, P., S.F. Khattak, and Z.J. Li, *Optimal and consistent protein glycosylation in mammalian cell culture*. Glycobiology, 2009. **19**(9): p. 936-949.
15. Schiestl, M., et al., *Acceptable changes in quality attributes of glycosylated biopharmaceuticals*. Nat Biotech, 2011. **29**(4): p. 310-312.
16. Sola, R.J. and K. Griebenow, *Effects of Glycosylation on the Stability of Protein Pharmaceuticals*. Journal of Pharmaceutical Sciences, 2009. **98**(4): p. 1223-1245.
17. Ly, M., T.N. Laremore, and R.J. Linhardt, *Proteoglycomics: Recent Progress and Future Challenges*. Omics-a Journal of Integrative Biology, 2010. **14**(4): p. 389-399.
18. Sasisekharan, R., R. Raman, and V. Prabhakar, *Glycomics Approach to Structure-Function Relationships of Glycosaminoglycans*. Annual Review of Biomedical Engineering, 2006. **8**(1): p. 181-231.
19. Yates, A., *Glycolipids and gliomas*. Molecular and Chemical Neuropathology, 1988. **8**(3): p. 157-180.
20. Brandenburg, K. and O. Holst, *Glycolipids: Distribution and Biological Function*, in *eLS*. 2001, John Wiley & Sons, Ltd.
21. Mayer, T.G., et al., *Synthesis Of Labeled Glycosyl Phosphatidyl Inositol (GPI) Anchors*. European Journal of Organic Chemistry, 1999. **1999**(10): p. 2563-2571.
22. Hooper, N.M., *GPI anchors*, in *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*. 2004, John Wiley & Sons, Ltd.

23. Furmanek, A. and J. Hofsteenge, *Protein C-mannosylation: Facts and questions*. Acta Biochimica Polonica, 2000. **47**(3): p. 781-789.
24. Kornfeld, R. and S. Kornfeld, *Assembly of Asparagine-Linked Oligosaccharides*. Annual Review of Biochemistry, 1985. **54**: p. 631-664.
25. Jenkins, N., R.B. Parekh, and D.C. James, *Getting the glycosylation right: Implications for the biotechnology industry*. Nat Biotech, 1996. **14**(8): p. 975-981.
26. Schwarz, F. and M. Aebi, *Mechanisms and principles of N-linked protein glycosylation*. Current Opinion in Structural Biology, 2011. **21**(5): p. 576-582.
27. Jones, J., S.S. Krag, and M.J. Betenbaugh, *Controlling N-linked glycan site occupancy*. Biochimica et Biophysica Acta (BBA) - General Subjects, 2005. **1726**(2): p. 121-137.
28. Durand, G. and N. Seta, *Protein Glycosylation and Diseases: Blood and Urinary Oligosaccharides as Markers for Diagnosis and Therapeutic Monitoring*. Clinical Chemistry, 2000. **46**(6): p. 795-805.
29. Yamashita, K., S. Hara-Kuge, and T. Ohkura, *Intracellular lectins associated with N-linked glycoprotein traffic*. Biochimica et Biophysica Acta (BBA) - General Subjects, 1999. **1473**(1): p. 147-160.
30. Hakomori, S., *Carbohydrate-to-carbohydrate interaction, through glycosynapse, as a basis of cell recognition and membrane organization*. Glycoconjugate Journal, 2004. **21**(3-4): p. 125-137.
31. Crean, S.M., et al., *N-linked sialylated sugar receptors support haematopoietic cell-osteoblast adhesions*. British Journal of Haematology, 2004. **124**(4): p. 534-546.
32. Bierhuizen, M.F.A., et al., *Structural Assessment of the N-linked Oligosaccharides of Cell-CAM 105 by Lectin-Agarose Affinity-Chromatography*. Glycoconjugate Journal, 1989. **6**(2): p. 195-208.
33. Berjonneau, C., et al., *Correlation between changes in cell adhesion and the ratio of N- to O-linked glycopeptides during chick embryo development*. Biology of the Cell, 1985. **52**(1): p. 21-26.
34. Byrne, B., G.G. Donohoe, and R. O'Kennedy, *Sialic acids: carbohydrate moieties that influence the biological and physical properties of biopharmaceutical proteins and living cells*. Drug Discovery Today, 2007. **12**(7-8): p. 319-326.
35. Hanson, S.R., et al., *The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(9): p. 3131-3136.
36. Morell, A.G., et al., *The Role of Sialic Acid in Determining the Survival of Glycoproteins in the Circulation*. Journal of Biological Chemistry, 1971. **246**(5): p. 1461-1467.
37. Cioffi, D.L., et al., *Terminal sialic acids are an important determinant of pulmonary endothelial barrier integrity*. American Journal of Physiology-Lung Cellular and Molecular Physiology, 2012. **302**(10): p. L1067-L1077.
38. Yamamoto, H., et al., *α ,6-Sialylation of Cell-Surface N-Glycans Inhibits Glioma Formation in Vivo*. Cancer Research, 2001. **61**(18): p. 6822-6829.
39. Seales, E.C., et al., *Hypersialylation of beta(1) integrins, observed in colon adenocarcinoma, may contribute to cancer progression by up-regulating cell motility*. Cancer Research, 2005. **65**(11): p. 4645-4652.
40. Hedlund, M., et al., *alpha 2-6-linked sialic acids on N-glycans modulate carcinoma differentiation in vivo*. Cancer Research, 2008. **68**(2): p. 388-394.
41. Axelsson, M.A., et al., *Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the O-glycosylation of mucins*. Glycobiology, 2001. **11**(8): p. 633-44.
42. Thorens, B. and P. Vassalli, *Chloroquine and ammonium chloride prevent terminal glycosylation of immunoglobulins in plasma cells without affecting secretion*. Nature, 1986. **321**(6070): p. 618-620.
43. Trombetta, E.S., *The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis*. Glycobiology, 2003. **13**(9): p. 77R-91R.
44. Ellgaard, L. and A. Helenius, *Quality control in the endoplasmic reticulum*. Nat Rev Mol Cell Biol, 2003. **4**(3): p. 181-191.

45. Spiro, R.G., *Role of N-linked polymannose oligosaccharides in targeting glycoproteins for endoplasmic reticulum-associated degradation*. Cellular and Molecular Life Sciences, 2004. **61**(9): p. 1025-1041.
46. Helenius, A. and M. Aebi, *Roles of N-Linked Glycans in the Endoplasmic Reticulum*. Annual Review of Biochemistry, 2004. **73**(1): p. 1019-1049.
47. Oliver, J.D., et al., *ERp57 Functions as a Subunit of Specific Complexes Formed with the ER Lectins Calreticulin and Calnexin*. Molecular Biology of the Cell, 1999. **10**(8): p. 2573-2582.
48. Stigliano, I.D., et al., *Glucosidase II and N-glycan mannose content regulate the half-lives of monoglucosylated species in vivo*. Molecular Biology of the Cell, 2011. **22**(11): p. 1810-1823.
49. Hammond, C., I. Braakman, and A. Helenius, *Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control*. Proceedings of the National Academy of Sciences, 1994. **91**(3): p. 913-917.
50. Okajima, T., A. Matsuura, and T. Matsuda, *Biological functions of glycosyltransferase genes involved in O-fucose glycan synthesis*. Journal of Biochemistry, 2008. **144**(1): p. 1-6.
51. Haltiwanger, R.S., et al., *O-Fucose and O-Glucose: Essential Modifications for Notch Function*. Glycobiology, 2008. **18**(11): p. 947-947.
52. Wells, L., S.A. Whalen, and G.W. Hart, *O-GlcNAc: a regulatory post-translational modification*. Biochemical and Biophysical Research Communications, 2003. **302**(3): p. 435-441.
53. Lommel, M. and S. Strahl, *Protein O-mannosylation: Conserved from bacteria to humans*. Glycobiology, 2009. **19**(8): p. 816-828.
54. Jensen, P.H., D. Kolarich, and N.H. Packer, *Mucin-type O-glycosylation – putting the pieces together*. Febs Journal, 2010. **277**(1): p. 81-94.
55. Marth, J.D., *Complexity in O-linked oligosaccharide biosynthesis engendered by multiple polypeptide N-acetylgalactosaminyltransferases*. Glycobiology, 1996. **6**(7): p. 701-705.
56. Porchet, N. and J.P. Aubert, *MUC genes: mucin or not mucin ? That is the question*. M S-Medecine Sciences, 2004. **20**(5): p. 569-574.
57. Steen, P.V.d., et al., *Concepts and Principles of O-Linked Glycosylation*. Critical Reviews in Biochemistry and Molecular Biology, 1998. **33**(3): p. 151-208.
58. Silverman, H.S., et al., *The contribution of tandem repeat number to the O-glycosylation of mucins*. Glycobiology, 2003. **13**(4): p. 265-277.
59. Lodish, H., et al., *Molecular cell biology; Third edition*. Molecular cell biology, Third edition. 1995: Scientific American Books Inc. {a}, 41 Madison Avenue, New York, New York 10010, USA. 1+1344p.
60. Henry, S., R. Oriol, and B. Samuelsson, *Lewis Histo-Blood Group System and Associated Secretory Phenotypes*. Vox Sanguinis, 1995. **69**(3): p. 166-182.
61. Lindén, S.K., et al., *Mucins in the mucosal barrier to infection*. Mucosal Immunology, 2008. **1**(3): p. 183-197.
62. Bansil, R. and B.S. Turner, *Mucin structure, aggregation, physiological functions and biomedical applications*. Current Opinion in Colloid & Interface Science, 2006. **11**(2-3): p. 164-170.
63. Gill, D.J., H. Clausen, and F. Bard, *Location, location, location: new insights into O-GalNAc protein glycosylation*. Trends in cell biology, 2011. **21**(3): p. 149-158.
64. Bennett, E.P., et al., *Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family*. Glycobiology, 2012. **22**(6): p. 736-756.
65. Springer, G.F., *Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy*. Journal of Molecular Medicine-Jmm, 1997. **75**(8): p. 594-602.
66. Yamada, T., et al., *Sialosyl Tn antigen expression is associated with the prognosis of patients with advanced gastric cancer*. Cancer, 1995. **76**(9): p. 1529-1536.
67. Yu, L.-G., *The oncofetal Thomsen–Friedenreich carbohydrate antigen in cancer progression*. Glycoconjugate Journal, 2007. **24**(8): p. 411-420.
68. Itzkowitz, S.H., et al., *Expression of Tn, Sialosyl-Tn, and T Antigens in Human Colon Cancer*. Cancer Research, 1989. **49**(1): p. 197-204.

69. Hedlund, M., et al., *Evidence for a human-specific mechanism for diet and antibody-mediated inflammation in carcinoma progression*. Proceedings of the National Academy of Sciences, 2008. **105**(48): p. 18936-18941.
70. McGuckin, M.A., et al., *Mucin dynamics and enteric pathogens*. Nature Reviews Microbiology, 2011. **9**(4): p. 265-278.
71. Marshall, B.J. and J.R. Warren, *Unidentified curved bacilli in the stomach of patients with gastritis and peptic-ulceration*. Lancet, 1984. **1**(8390): p. 1311-1315.
72. Polk, D.B. and R.M. Peek, *Helicobacter Pylori: gastric cancer and beyond*. Nature Reviews Cancer, 2010. **10**(8): p. 593-593.
73. Kobayashi, M., et al., *Roles of gastric mucin-type O-glycans in the pathogenesis of Helicobacter pylori infection*. Glycobiology, 2009. **19**(5): p. 453-461.
74. Farinha, P. and R.D. Gascoyne, *Helicobacter pylori and MALT Lymphoma*. Gastroenterology, 2005. **128**(6): p. 1579-1605.
75. Ilver, D., et al., *Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging*. Science, 1998. **279**(5349): p. 373-377.
76. Bäckström, A., et al., *Metastability of Helicobacter pylori bab adhesin genes and dynamics in Lewis b antigen binding*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(48): p. 16923-16928.
77. Mahdavi, J., et al., *Helicobacter pylori SabA Adhesin in Persistent Infection and Chronic Inflammation*. Science, 2002. **297**(5581): p. 573-578.
78. Lindén, S., et al., *Strain- and blood group-dependent binding of Helicobacter pylori to human gastric MUC5AC glycoforms*. Gastroenterology, 2002. **123**(6): p. 1923-1930.
79. Styer, C.M., et al., *Expression of the BabA Adhesin during Experimental Infection with Helicobacter pylori*. Infection and Immunity, 2010. **78**(4): p. 1593-1600.
80. Kawakubo, M., et al., *Natural antibiotic function of a human gastric mucin against Helicobacter pylori infection*. Science, 2004. **305**(5686): p. 1003-6.
81. Honke, K. and N. Taniguchi, *Sulfotransferases and sulfated oligosaccharides*. Medicinal Research Reviews, 2002. **22**(6): p. 637-654.
82. Torii, T., M. Fukuta, and O. Habuchi, *Sulfation of sialyl N-acetyllactosamine oligosaccharides and fetuin oligosaccharides by keratan sulfate Gal-6-sulfotransferase*. Glycobiology, 2000. **10**(2): p. 203-211.
83. Brockhausen, I., *Sulphotransferases acting on mucin-type oligosaccharides*. Biochemical Society Transactions, 2003. **31**(2): p. 318-325.
84. Estrella, R.P., et al., *Graphitized Carbon LC-MS Characterization of the Chondroitin Sulfate Oligosaccharides of Aggrecan*. Analytical Chemistry, 2007. **79**(10): p. 3597-3606.
85. Toyama-Sorimachi, N., et al., *Widespread Expression of Chondroitin Sulfate-type Serylcins with CD44 Binding Ability in Hematopoietic Cells*. Journal of Biological Chemistry, 1997. **272**(42): p. 26714-26719.
86. Albertini, R., et al., *Chondroitin-4-sulfate protects high-density lipoprotein against copper-dependent oxidation*. Archives of Biochemistry and Biophysics, 1999. **365**(1): p. 143-149.
87. Shen, Y.J., et al., *PTP sigma Is a Receptor for Chondroitin Sulfate Proteoglycan, an Inhibitor of Neural Regeneration*. Science, 2009. **326**(5952): p. 592-596.
88. Mulloy, B. and M.J. Forster, *Conformation and dynamics of heparin and heparan sulfate*. Glycobiology, 2000. **10**(11): p. 1147-1156.
89. Yamashita, Y., et al., *Effect of heparin on pulmonary fibroblasts and vascular cells*. Thorax, 1992. **47**(8): p. 634-639.
90. Xia, C.-Q. and K.-J. Kao, *Heparin Induces Differentiation of CD1a+ Dendritic Cells from Monocytes: Phenotypic and Functional Characterization*. The Journal of Immunology, 2002. **168**(3): p. 1131-1138.
91. Maldonado, B.A. and L.T. Furcht, *Involvement of integrins with adhesion-promoting, heparin-binding peptides of type IV collagen in cultured human corneal epithelial cells*. Investigative Ophthalmology & Visual Science, 1995. **36**(2): p. 364-72.
92. Kovensky, J., *Sulfated Oligosaccharides: New Targets for Drug Development? Current Medicinal Chemistry*, 2009. **16**(18): p. 2338-2344.

93. Fukuta, M., et al., *Molecular cloning and characterization of human keratan sulfate Gal-6-sulfotransferase*. Journal of Biological Chemistry, 1997. **272**(51): p. 32321-32328.
94. Hooper, L.V., S.M. Manzella, and J.U. Baenziger, *From legumes to leukocytes: biological roles for sulfated carbohydrates*. FASEB J, 1996. **10**(10): p. 1137-46.
95. Morita, I., et al., *Expression and function of the HNK-1 carbohydrate*. Journal of Biochemistry, 2008. **143**(6): p. 719-724.
96. Kawashima, H., *Roles of sulfated Glycans in lymphocyte homing*. Biological & Pharmaceutical Bulletin, 2006. **29**(12): p. 2343-2349.
97. Davril, M., et al., *The sialylation of bronchial mucins secreted by patients suffering from cystic fibrosis or from chronic bronchitis is related to the severity of airway infection*. Glycobiology, 1999. **9**(3): p. 311-321.
98. Schulz, B.L., et al., *Mucin glycosylation changes in cystic fibrosis lung disease are not manifest in submucosal gland secretions*. Biochemical Journal, 2005. **387**(Pt 3): p. 911-919.
99. Dell, A. and H.R. Morris, *Glycoprotein structure determination mass spectrometry*. Science, 2001. **291**(5512): p. 2351-2356.
100. Mechref, Y. and M.V. Novotny, *Structural Investigations of Glycoconjugates at High Sensitivity*. Chemical Reviews, 2002. **102**(2): p. 321-370.
101. Harvey, D.J., *Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates*. International Journal of Mass Spectrometry, 2003. **226**(1): p. 1-35.
102. Karlsson, N.G., et al., *Use of graphitised carbon negative ion LC-MS to analyse enzymatically digested glycosaminoglycans*. J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **824**(1-2): p. 139-47.
103. Desantos-Garcia, J.L., et al., *Enhanced sensitivity of LC-MS analysis of permethylated N-glycans through online purification*. Electrophoresis, 2011. **32**(24): p. 3516-3525.
104. Marino, K., et al., *Method for milk oligosaccharide profiling by 2-aminobenzamide labeling and hydrophilic interaction chromatography*. Glycobiology, 2011. **21**(10): p. 1317-30.
105. Wuhrer, M., A.R. de Boer, and A.M. Deelder, *Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry*. Mass Spectrom Rev, 2009. **28**(2): p. 192-206.
106. Wuhrer, M., C.A. Koeleman, and A.M. Deelder, *Two-dimensional HPLC separation with reverse-phase-nano-LC-MS/MS for the characterization of glycan pools after labeling with 2-aminobenzamide*. Methods Mol Biol, 2009. **534**: p. 79-91.
107. Zamfir, A., et al., *Structural investigation of chondroitin/dermatan sulfate oligosaccharides from human skin fibroblast decorin*. Glycobiology, 2003. **13**(11): p. 733-742.
108. Wuhrer, M., A.M. Deelder, and C.H. Hokke, *Protein glycosylation analysis by liquid chromatography-mass spectrometry*. Journal of Chromatography B, 2005. **825**(2): p. 124-133.
109. Novotny, M.V. and Y. Mechref, *New hyphenated methodologies in high-sensitivity glycoprotein analysis*. Journal of Separation Science, 2005. **28**(15): p. 1956-1968.
110. Hennion, M.-C., *Graphitized carbons for solid-phase extraction*. Journal of Chromatography A, 2000. **885**(1-2): p. 73-95.
111. Ruhaak, L.R., A.M. Deelder, and M. Wuhrer, *Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry*. Analytical and Bioanalytical Chemistry, 2009. **394**(1): p. 163-74.
112. Thomsson, K.A., et al., *Enhanced detection of sialylated and sulfated glycans with negative ion mode nanoliquid chromatography/mass spectrometry at high pH*. Anal Chem, 2010. **82**(4): p. 1470-7.
113. Budai, L., et al., *Analysis of complex oligosaccharides using graphitized carbon liquid chromatography/mass spectrometry*. European Journal of Mass Spectrometry, 2008. **14**(6): p. 419-422.
114. Davies, M.J., et al., *Use of a porous graphitised carbon column for the high-performance liquid chromatography of oligosaccharides, alditols and glycopeptides with subsequent mass spectrometry analysis*. Journal of Chromatography A, 1993. **646**(2): p. 317-326.
115. Nwosu, C.C., et al., *Comparison of the Human and Bovine Milk N-Glycome via High-Performance Microfluidic Chip Liquid Chromatography and Tandem Mass Spectrometry*. Journal of Proteome Research, 2012. **11**(5): p. 2912-2924.

116. Karlsson, N.G., et al., *Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis*. *Rapid Commun Mass Spectrom*, 2004. **18**(19): p. 2282-92.
117. Doohan, R., et al., *Negative Ion CID Fragmentation of O-linked Oligosaccharide Aldoses—Charge Induced and Charge Remote Fragmentation*. *Journal of The American Society for Mass Spectrometry*, 2011. **22**(6): p. 1052-1062.
118. Schulz, B.L., N.H. Packer, and N.G. Karlsson, *Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis*. *Analytical Chemistry*, 2002. **74**(23): p. 6088-6097.
119. Barroso, B., M. Didraga, and R. Bischoff, *Analysis of proteoglycans derived sulphated disaccharides by liquid chromatography/mass spectrometry*. *Journal of Chromatography A*, 2005. **1080**(1): p. 43-48.
120. Kenny, D., et al., *Perspective and Review of Mass Spectrometric Based Sulfoglycomics of N-Linked and O-Linked Oligosaccharides*. *Current Proteomics*, 2011. **8**(4): p. 278-296.
121. Thomsson, K.A., et al., *Enhanced Detection of Sialylated and Sulfated Glycans with Negative Ion Mode Nanoliquid Chromatography/Mass Spectrometry at High pH*. *Analytical Chemistry*, 2010. **82**(4): p. 1470-1477.
122. Delaney, J. and P. Vouros, *Liquid chromatography ion trap mass spectrometric analysis of oligosaccharides using permethylated derivatives*. *Rapid Communications in Mass Spectrometry*, 2001. **15**(5): p. 325-334.
123. Songsrirote, K., et al., *Development and application of mass spectrometric methods for the analysis of progranulin N-glycosylation*. *Journal of Proteomics*, 2010. **73**(8): p. 1479-1490.
124. Fenn, J., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. *Science*, 1989. **246**(4926): p. 64-71.
125. Gaskell, S.J., *Electrospray: Principles and practice*. *Journal of Mass Spectrometry*, 1997. **32**(7): p. 677-688.
126. Bakhoun, S.F.W. and G.R. Agnes, *Study of chemistry in droplets with net charge before and after Coulomb explosion: Ion-induced nucleation in solution and implications for ion production in an electrospray*. *Analytical Chemistry*, 2005. **77**(10): p. 3189-3197.
127. Zenobi, R. and R. Knochenmuss, *Ion formation in MALDI mass spectrometry*. *Mass Spectrometry Reviews*, 1998. **17**(5): p. 337-366.
128. Douglas, D.J., A.J. Frank, and D.M. Mao, *Linear ion traps in mass spectrometry*. *Mass Spectrometry Reviews*, 2005. **24**(1): p. 1-29.
129. Makarov, A., *Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis*. *Analytical Chemistry*, 2000. **72**(6): p. 1156-1162.
130. Penn, S.G., M.T. Cancilla, and C.B. Lebrilla, *Collision-induced dissociation of branched oligosaccharide ions with analysis and calculation of relative dissociation thresholds*. *Anal Chem*, 1996. **68**(14): p. 2331-9.
131. Zaia, J. and C.E. Costello, *Tandem mass spectrometry of sulfated heparin-like glycosaminoglycan oligosaccharides*. *Anal Chem*, 2003. **75**(10): p. 2445-55.
132. Wuhrer, M., et al., *Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum*. *Proteomics*, 2007. **7**(22): p. 4070-4081.
133. Kenny, D.T., et al., *Presence of Terminal N-acetylgalactosamine β 1-4N-acetylglucosamine Residues on O-linked Oligosaccharides from Gastric MUC5AC: Involvement in Helicobacter pylori Colonization?* *Glycobiology*, 2012.
134. Karlsson, N.G., B.L. Schulz, and N.H. Packer, *Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn*. *J Am Soc Mass Spectrom*, 2004. **15**(5): p. 659-72.
135. Kenny, D.T., S.M.A. Issa, and N.G. Karlsson, *Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation*. *Rapid Communications in Mass Spectrometry*, 2011. **25**(18): p. 2611-2618.
136. Wuhrer, M., et al., *Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments*. *Rapid Communications in Mass Spectrometry*, 2006. **20**(11): p. 1747-1754.

137. Harvey, D.J., et al., *"Internal residue loss": rearrangements occurring during the fragmentation of carbohydrates derivatized at the reducing terminus*. Analytical Chemistry, 2002. **74**(4): p. 734-740.
138. Jedrychowski, M.P., et al., *Evaluation of HCD- and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics*. Molecular & Cellular Proteomics, 2011.
139. Olsen, J.V., et al., *Higher-energy C-trap dissociation for peptide modification analysis*. Nat Meth, 2007. **4**(9): p. 709-712.
140. 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.
141. Cooper, C.A., E. Gasteiger, and N.H. Packer, *GlycoMod – A software tool for determining glycosylation compositions from mass spectrometric data*. PROTEOMICS, 2001. **1**(2): p. 340-349.
142. Ceroni, A., et al., *GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans†*. Journal of Proteome Research, 2008. **7**(4): p. 1650-1659.
143. Ceroni, A., A. Dell, and S.M. Haslam, *The GlycanBuilder: a fast, intuitive and flexible software tool for building and displaying glycan structures*. Source code for biology and medicine, 2007. **2**: p. 3.
144. Hayes, C.A., et al., *UniCarb-DB: a database resource for glycomic discovery*. Bioinformatics, 2011. **27**(9): p. 1343-1344.
145. Mariño, K., et al., *A systematic approach to protein glycosylation analysis: a path through the maze*. Nature Chemical Biology, 2010. **6**(10): p. 713-723.
146. Hashimoto, K., et al., *KEGG as a glycome informatics resource*. Glycobiology, 2006. **16**(5): p. 63R-70.
147. Campbell, M.P., et al., *GlycoBase and autoGU: tools for HPLC-based glycan analysis*. Bioinformatics, 2008. **24**(9): p. 1214-1216.
148. Harvey, D.J., *Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry*. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2011. **879**(17-18): p. 1196-1225.
149. Kang, P., et al., *Solid-phase permethylation of glycans for mass spectrometric analysis*. Rapid Communications in Mass Spectrometry, 2005. **19**(23): p. 3421-3428.
150. Ciucanu, I. and F. Kerek, *A simple and rapid method for the permethylation of carbohydrates*. Carbohydrate Research, 1984. **131**(2): p. 209-217.
151. Hakomori, S., *A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide*. Journal of Biochemistry, 1964. **55**: p. 205-208.
152. Yoo, E. and I. Yoon, *Applications of tandem mass spectrometry in the structure determination of permethylated sialic acid-containing oligosaccharides*. Bulletin of the Korean Chemical Society, 2005. **26**(9): p. 1347-1353.
153. Karlsson, N.G., et al., *Glycosylation differences between pig gastric mucin populations: A comparative study of the neutral oligosaccharides using mass spectrometry*. Biochemical Journal, 1997. **326**: p. 911-917.
154. Hoja-Lukowicz, D., et al., *High-mannose-type oligosaccharides from human placental arylsulfatase A are core fucosylated as confirmed by MALDI MS*. Glycobiology, 2000. **10**(6): p. 551-557.
155. Zaia, J., et al., *Structural analysis of cartilage proteoglycans and glycoproteins using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*. Anal Biochem, 2000. **277**(1): p. 94-103.
156. Liedtke, S., et al., *Characterization of N-glycans from mouse brain neural cell adhesion molecule*. Glycobiology, 2001. **11**(5): p. 373-84.
157. Wilson, N.L., et al., *Sequential analysis of N- and O-linked glycosylation of 2D-PAGE separated glycoproteins*. J Proteome Res, 2002. **1**(6): p. 521-9.
158. Ritchie, G., et al., *Identification of N-linked carbohydrates from severe acute respiratory syndrome (SARS) spike glycoprotein*. Virology, 2010. **399**(2): p. 257-69.

159. Tissot, B., et al., *Glycoproteomics: Past, present and future*. Febs Letters, 2009. **583**(11): p. 1728-1735.
160. Rebecchi, K.R., C.L. Woodin, and H. Desaire, *Recent Mass Spectrometric Based Methods in Quantitative N-linked Glycoproteomics*. Current Proteomics, 2011. **8**(4): p. 269-277.
161. Wuhrer, M., et al., *Glycoproteomics based on tandem mass spectrometry of glycopeptides*. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, 2007. **849**(1-2): p. 115-128.
162. Morelle, W. and J.-C. Michalski, *Analysis of protein glycosylation by mass spectrometry*. Nat. Protocols, 2007. **2**(7): p. 1585-1602.
163. Yu, S.-Y., et al., *Enabling techniques and strategic workflow for sulfoglycomics based on mass spectrometry mapping and sequencing of permethylated sulfated glycans*. Glycobiology, 2009. **19**(10): p. 1136-1149.
164. Kang, P., Y. Mechref, and M.V. Novotny, *High-throughput solid-phase permethylation of glycans prior to mass spectrometry*. Rapid Communications in Mass Spectrometry, 2008. **22**(5): p. 721-734.
165. Price, N.P.J., *Permethylation linkage analysis techniques for residual carbohydrates*. Applied Biochemistry and Biotechnology, 2008. **148**(1-3): p. 271-276.
166. Lam, M.P.Y., et al., *Online combination of reversed-phase/reversed-phase and porous graphitic carbon liquid chromatography for multicomponent separation of proteomics and glycoproteomics samples*. 2011. **32**(21): p. 2930-2940.
167. Christiansen, M.N., et al., *Challenges of Determining O-Glycopeptide Heterogeneity: A Fungal Glucanase Model System*. Analytical Chemistry, 2010. **82**(9): p. 3500-3509.
168. Haslam, S.M., S.J. North, and A. Dell, *Mass spectrometric analysis of N- and O-glycosylation of tissues and cells*. Current Opinion in Structural Biology, 2006. **16**(5): p. 584-591.
169. Pan, S., et al., *Mass Spectrometry Based Glycoproteomics—From a Proteomics Perspective*. Molecular & Cellular Proteomics, 2011. **10**(1).
170. Norris, G.E., et al., *Purification and crystallisation of the endoglycosidase PNGase-F, a peptide-N-glycosidase from Flavobacterium meningosepticum*. Journal of Molecular Biology, 1994. **241**(4): p. 624-626.
171. Altmann, F., et al., *Characterisation of peptide-N-4-(N-acetyl-beta-glucosaminyl)asparagine amidase A and its N-glycans*. European Journal of Biochemistry, 1998. **252**(1): p. 118-123.
172. Chiba, A., et al., *Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystroglycan The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin*. Journal of Biological Chemistry, 1997. **272**(4): p. 2156-2162.
173. Lei, M., M.V. Novotny, and Y. Mechref, *Sequential enrichment of sulfated glycans by strong anion-exchange chromatography prior to mass spectrometric measurements*. J. Am. Soc. Mass Spectrom., 2010. **21**(3): p. 348-357.
174. Garenaux, E., et al., *A single step method for purification of sulfated oligosaccharides*. Glycoconjugate Journal, 2008. **25**(9): p. 903-915.
175. Toyoda, M., H. Narimatsu, and A. Kameyama, *Enrichment Method of Sulfated Glycopeptides by a Sulfate Emerging and Ion Exchange Chromatography*. Analytical Chemistry, 2009. **81**(15): p. 6140-6147.
176. Hebert, D.N., S.C. Garman, and M. Molinari, *The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags*. Trends in cell biology, 2005. **15**(7): p. 364-370.
177. Xia, B., et al., *Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis*. Glycobiology, 2005. **15**(8): p. 747-775.
178. Rudd, P.M. and R.A. Dwek, *Rapid, sensitive sequencing of oligosaccharides from glycoproteins*. Current Opinion in Biotechnology, 1997. **8**(4): p. 488-497.
179. Thomsson, K.A., et al., *MUC5B glycosylation in human saliva reflects blood group and secretor status*. Glycobiology, 2005. **15**(8): p. 791-804.
180. Zaia, J., et al., *The role of mobile protons in negative ion CID of oligosaccharides*. J Am Soc Mass Spectrom, 2007. **18**(5): p. 952-60.

181. Balderrama, G.D., et al., *Analysis of sulfated peptides from the skin secretion of the *Pachymedusa dactylosa* frog using IMAC-Ga enrichment and high-resolution mass spectrometry*. Rapid Communications in Mass Spectrometry, 2011. **25**(8): p. 1017-1027.
182. Molloy, M.P., et al., *Proteomic analysis of the *Escherichia coli* outer membrane*. European Journal of Biochemistry, 2000. **267**(10): p. 2871-2881.
183. Varki, A., *Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells*. Cold Spring Harbor perspectives in biology, 2011. **3**(6).
184. Hayes, C., et al., *Glyco-Bioinformatic and Statistical Analysis of Inflammatory Response in Different Tissue Types*. Glycobiology, 2010. **20**(11): p. 1469-1470.
185. Lindén, S.K., et al., *Four Modes of Adhesion are Used During *Helicobacter pylori* Binding to Human Mucins in the Oral and Gastric Niches*. Helicobacter, 2008. **13**(2): p. 81-93.
186. Newman, W. and E.A. Kabat, *Immunochemical studies on blood groups. Structures and immunochemical properties of nine oligosaccharides from B-active and non-B-active blood group substances of horse gastric mucosae*. Arch Biochem Biophys, 1976. **172**(2): p. 353-50.
187. Ikehara, Y., et al., *Apical Golgi localization of *N,N'*-diacetyllactosylamine synthase, *beta*4GalNAc-T3, is responsible for LacdiNAc expression on gastric mucosa*. Glycobiology, 2006. **16**(9): p. 777-85.
188. Toyoda, M., et al., *Quantitative Derivatization of Sialic Acids for the Detection of Sialoglycans by MALDI MS*. Analytical Chemistry, 2008. **80**(13): p. 5211-5218.
189. Reis, C.A., et al., *Alterations in glycosylation as biomarkers for cancer detection*. Journal of Clinical Pathology, 2010. **63**(4): p. 322-329.
190. Pilatte, Y., J. Bignon, and C.R. Lambre, *Sialic acids As important molecules in the regulation of the immune-system - Pathophysiological implications of sialidases in immunity*. Glycobiology, 1993. **3**(3): p. 201-218.
191. Chen, P.F. and S.W. Harcum, *Effects of amino acid additions on ammonium stressed CHO cells*. Journal of Biotechnology, 2005. **117**(3): p. 277-286.
192. Moreland, L., G. Bate, and P. Kirkpatrick, *Abatacept*. Nat Rev Drug Discov, 2006. **5**(3): p. 185-186.
193. Kim, H.J. and H.-J. Kim, *The glycosylation and pharmacokinetics of CTLA4Ig produced in rice cells*. Biological & Pharmaceutical Bulletin, 2007. **30**(10): p. 1913-1917.
194. Senger, R.S. and M.N. Karim, *Effect of Shear Stress on Intrinsic CHO Culture State and Glycosylation of Recombinant Tissue-Type Plasminogen Activator Protein*. Biotechnology Progress, 2003. **19**(4): p. 1199-1209.
195. Nyberg, G.B., et al., *Metabolic effects on recombinant interferon-gamma glycosylation in continuous culture of Chinese hamster ovary cells*. Biotechnology and Bioengineering, 1999. **62**(3): p. 336-347.
196. Kenny, D.T., et al., *Glycomic Analysis of Membrane-Associated Proteins*, in *Sample Preparation in Biological Mass Spectrometry*, A.R. Ivanov and A.V. Lazarev, Editors. 2011, Springer. p. 498-513.

9 Personal Involvement in Each Paper:

Paper I: It is primarily a review article. I and my supervisor Dr. Niclas Karlsson were the major contributors to the article. I did the background research. It was a collaborative effort whereby I would provide text to Dr. Karlsson and he would provide his input. One collaborator aided in the preparation of figure 2 and provided figure 7B and the other collaborator provided figure 3.

Paper II: I performed the major portion of the experimental work. My collaborator, Samah Issa, was involved in the HCD fragmentation of sulfated and sialyl-sulfated oligosaccharide standards. All analysis was performed by me. I wrote the article in co-operation with Dr. Karlsson.

Paper III: I, Liaqat Ali and Dr Karlsson contributed equally. I was primarily involved in developing and writing the protocol for the preparation and analysis by LC-MS of membrane associated oligosaccharides and I generated the data from the CHO membrane sample.

Paper IV: I performed all the experimental work and analysis of the results. All data was generated by me. I wrote the article with assistance from Dr. Karlsson. My collaborator Kristina Thomsson provided technical assistance while analysing the samples by LC-MS.

Paper V: I performed the major portion of the experimental work. My collaborators Emma Skoog performed the binding and proliferation studies and Weston Struwe provided technical assistance with the preparation of the exo-hexosaminidase digestion. All data, apart from the binding and proliferation assays, was generated by me. I wrote the article in co-operation with Dr. Karlsson. Additional input regarding the biology of *H. pylori* was provided by our collaborator Emma Skoog..

All additional data in the discussion of the thesis was generated by me.

Perspective and Review of Mass Spectrometric Based Sulfoglycomics of *N*-Linked and *O*-Linked Oligosaccharides

Diarmuid Kenny^{1,2}, Catherine A. Hayes², Chunseng Jin² and Niclas G. Karlsson^{2,3}

¹*School of Chemistry, National University of Ireland, Galway, Ireland* ²*Medical Biochemistry, University of Gothenburg, 405 30 Gothenburg, Sweden*

Abstract: The research area of glycomics is maturing and it is realized that it is as difficult to identify the relevance of glycosylation amongst biological house-keeping as it is with any of the other omics techniques. This review focuses on sulfoglycomics of *N*-linked and *O*-linked oligosaccharides to identify relevant biological glycomics questions. The review describes how sulfated oligosaccharides have been analyzed in the past, with the focus on MS and its current application using ESI and MALDI. Also included are the particular issues of sample preparation of sulfated oligosaccharides, including enrichment and separation using capillary and nano liquid chromatography. Structural characterization of sulfated oligosaccharides is currently carried out by collision induced dissociation, but recent developments focus on increasing the structural information by improved derivatization techniques followed by MS analysis and fragmentation of derivatized intact sulfooligosaccharides. Significant progress has been shown in this field in the application of permethylation derivatization. Current trends for fragmentation of biomolecules using novel techniques are also discussed with emphasis on sulfoglycomics. The conclusion of the review is that in order to understand the largely unexplored area of sulfation of oligosaccharides, there are techniques that need to be adapted and optimized to address the structural characterization, ranging all the way from appropriate sample preparation to effective MS analysis.

Keywords: Glycomics, glycosylation, mass spectrometry, post translational modifications, sulfate.

Introduction

With the finding that the limited number of genes in the human genome would not encode for the complexity of proteins that could explain the diverse functions of the cell, there is an acceptance that life science has to go beyond traditional genomics and proteomics to find the answers. The increased understanding of post-translational modifications has led to adopting the “one gene-many functions” concept to explain some of the deficiencies of a limited genome. The important role of phosphorylation as a post-translational modification is undisputed, and as a regulator of intracellular signaling its importance is obvious. However, in a complex multicellular organism there is also an obvious need for extracellular signaling. This review will discuss the analysis of glycosylation by mass spectrometry (MS). MS can be used to address how oligosaccharides are modified by sulfation to change the way glycoproteins “signal to” and interacting with other biomolecules. Sulfation is widely known in the area of proteoglycans with their high molecular mass glycosaminoglycans; heparin/heparan sulfate, chondroitin/dermatan sulfate, and keratan sulfate. For these molecules consistent methods using negative ion MS are currently being developed in various glycomic facilities throughout the world [1-3] and have been reviewed elsewhere. The sulfation of *N*-linked and *O*-linked oligosaccharides is the focus of this review. Examples of the biological role of *N*-linked and *O*-linked oligosaccharide sulfation (Fig. 1) are scarce and are in stark contrast to its ubiquitous expression. This is probably a reflection of the lack of analytical tools to investigate the sulfation rather than the lack of biological importance. Sulfation of *N*-linked oligosaccharides has been shown to be involved in the regulation of lutenizing hormones[4] and for activation of human natural killer cells (the sulfated HNK-epitope) [5].

³To whom correspondence should be addressed: Niclas Karlsson, Medical Biochemistry, University of Gothenburg, P.O. Box 440, 405 30 Gothenburg, Sweden. +46 31 786 6528. Fax +46 31 41 6108. E-mail: niclas.karlsson@medkem.gu.se

The sulfation of low molecular mass *N*-linked and *O*-linked *N*-acetyllactosamine containing oligosaccharides is structurally similar to the polyactosamine sulfation of keratan sulfate. With more sophisticated analytical techniques being developed, the larger sulfated *N*-linked and *O*-linked oligosaccharides will be continuous with keratan sulfate (KS) oligosaccharides, with the sulfation of *N*-acetyllactosamine as the common denominator. Many of the sulfotransferases have been shown to be the same for the smaller *N*- and *O*-linked oligosaccharides and the large polymerized keratan sulfate oligosaccharides[6]. For instance, it is likely that the definition of mucin type sulfation and KS will be based on their localization, the extracellular matrix for KS and mucosal surfaces for mucins rather than actual molecular differences.

Sulfation of *O*-linked oligosaccharides is most famously known in the homing of lymphocytes to peripheral lymph nodes mediated by sulfated oligosaccharides on special endothelium called high endothelial venules[7]. The initial binding of the lymphocytes is via a membrane bound C-type lectin, L-selectin, to the oligosaccharide epitope 6-sulfo sialyl Lewis x[8]. The isomeric sulfated epitope 6'-sulfo sialyl Lewis x, a recently discovered sulfated ligand, has been shown to be specific for human siglec-8 expressed on eosinophiles [9], but it is not known if its natural oligosaccharide ligand is expressed on *N*-linked or *O*-linked oligosaccharides or on other types on glycoconjugates. Sulfation of *O*-linked oligosaccharides has been the topic of discussion for mucin glycosylation in the respiratory tract [10-12], where it was believed to be involved in the pathology of cystic fibrosis. However, recent discoveries have shown that the alteration of respiratory mucin oligosaccharides is related to secondary effects caused by pulmonary bacterial infections and inflammation[13-15].

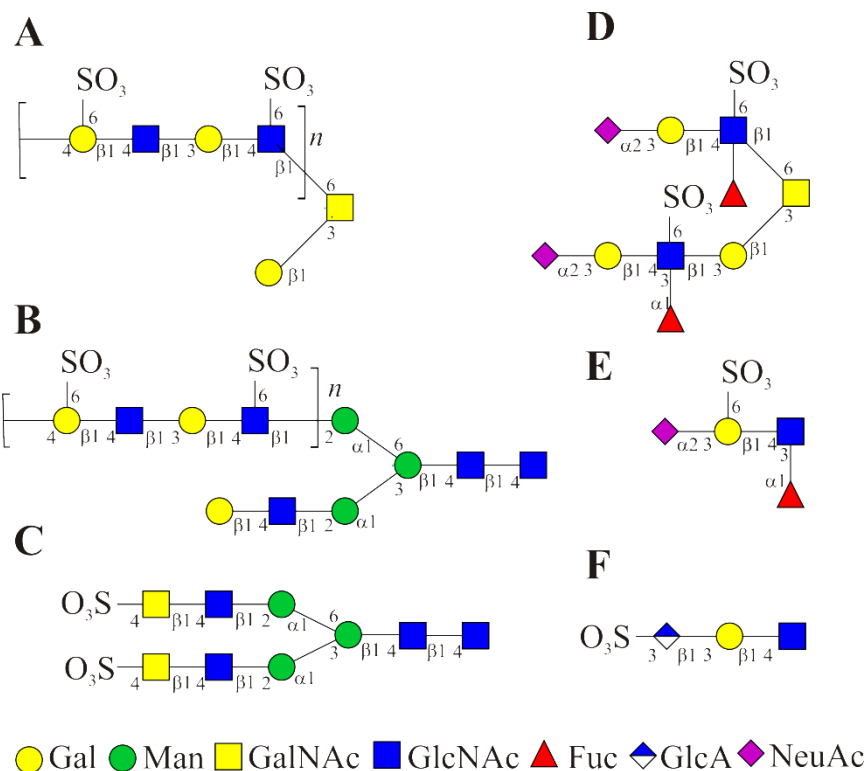


Fig (1) Biological sulfated epitopes found on *N*-linked and *O*-linked oligosaccharides. Examples of typical *O*-linked keratan sulfate (A) and *N*-linked keratan sulfate (B) with 6 linked sulfate, *N*-linked oligosaccharides from lutenizing hormones with 4 linked sulfate (C), *O*-linked oligosaccharide with 6 sulfo sialyl Lewis x as the ligand for L-selectin (D) and the oligosaccharide epitopes 6' sulfo sialyl Lewis x, which is the ligand for siglec 8 (E) and the sulfated HNK-epitope (F)

Historically, the methods for detection and characterization of sulfated oligosaccharides relied on a plethora of chemical and analytical methods. Radioactive labeling is still one of the most utilized methods to detect sulfoglycoproteins. Chemical methods for characterization involved monosaccharide composition analysis, sulfate analysis, chemical desulfation, linkage analysis by partially methylated alditol acetates and MS using FAB-MS with or without derivatization. The accuracy of the structural analysis was further improved by large scale isolation followed by $^1\text{H-NMR}$ [12, 16]. The increased sensitivity of MS using ESI and MALDI meant that

the methods for the analysis of sulfated oligosaccharides in many cases were miniaturized in order to adopt it to the omic era.

How to address sulfoglycomics?

As was described in the previous section, the importance of sulfation has been shown in endocrinology and immunology and this is likely to increase as the analytical techniques improve and new experimental platforms become available. The modern approach to sulfoglycomics will of course be inspired by the early pioneers in this field. It is likely that sulfoglycomics will evolve from existing experimental platforms for glycomics [17, 18]. Modern sulfoglycomics approaches would be based on standard derivatization method, separation techniques and analyses, so that the researchers in the field feel comfortable in adopting the platform in the context of their own research.

MS is often the technique of choice for the analysis of oligosaccharides including sulfated oligosaccharides. Based on previous historical examples and current state of the art analytical approaches (Table 1) a consensus platform for sulfoglycomics can be proposed (Fig 2). MS is performed on either native or derivatized sulfated oligosaccharides after they have been released and enriched from a single glycoprotein or from mixtures of enriched sulfoglycoproteins. Often the analysis of the native sulfated oligosaccharides is performed in negative ion mode as the sulfate group naturally imparts a negative charge that aids ionization (Table 2).

Table 1. Summary of approaches used for MS analysis of sulfated *N*-linked and *O*-linked oligosaccharides

Chromatography	MS	Mode	Sample Type	Adducts	MS ^{n a}
PGC	ESI	+	Free, Alditols	H ⁺ , Na ⁺ Other Metal Ions	+
			Permethylated	Na ⁺ Other Metal Ions	++
			Reducing end	Derivatives of quaternary amines and amines with high pKa gives mixed H ⁺ /metal ion adducts, other gives preferentially metal ion adducts	+ / ++
Reversed Phase		-	Free, Alditols	[M-xH] ^{x-} ions	+
			Reducing End	Negatively charged derivatives give [M-xH] ^{x-} ions	+
			Free, Alditols	H ⁺ , Na ⁺ Other Metal Ion's	+
Normal Phase	MALDI	+	Permethylated Peracetylated	Na ⁺ Other Metal Ion's	++
			Reducing End	Derivatives of quaternary amines and amines with high pKa gives H ⁺ adducts, other gives preferentially metal adducts	+ / ++
			Free, Alditols	[M-xH] ^{x-} ions of both neutral and acidic oligosaccharides, also Cl ⁻ adducts for neutral	+
HILIC		-	Permethylated	[M-xH] ^{x-} ions of both neutral and acidic oligosaccharides	+
			Reducing End	[M-xH] ^{x-} ions of both neutral and acidic oligosaccharides	+

^a+ means good, ++ means excellent

Analysis in positive ion mode requires derivatization of the oligosaccharide to impart a net positive charge or by forming adducts with positively charged metal ions. In general for glycomics, derivatization of oligosaccharides can be performed in order to improve structural and positional information by fragmentation MSⁿ[19]. This will also be true for sulfated oligosaccharides. Derivatization can be performed either on the reducing end or as an overall modification technique such as peracetylation or permethylation. Peracetylation has been performed on a regular basis for sulfated oligosaccharides [20-22], whereas permethylation has been less widely used since it can be problematic due to the labile nature of the sulfate group, which makes the work up procedure after derivatization difficult.

While most of the early work with sulfated oligosaccharides has been done using FAB-MS, both ESI and MALDI are now used preferentially for the analysis of sulfated oligosaccharide (Table 2). ESI appears to be a softer ionisation process for sulfated oligosaccharides, whereas dissociation of the sulfate group can be a challenging obstacle with MALDI. As a consequence, a certain amount of trial and error is required to determine the optimal MALDI methodology that will suppress the dissociation of the sulfates while still offering good sensitivity.

Although MS is the key for a successful sulfoglycomic research initiative, the additional pieces of the sulfoglycomic platform are also important. Enrichment of sulfoglycoproteins should be initiated prior to the release of sulfated oligosaccharides. This can be based on charge (e.g. anion exchange chromatography)[23] or by size/density (e.g. techniques used for mucin isolation). One of the most efficient methods for the isolation of glycoproteins is using 1D or 2D SDS-PAGE. This has been used to isolate individual sulfoglycoproteins with subsequent release and identification of sulfated oligosaccharides [24-26].

Once a glycomic sample has been identified, the next step in the analysis involves the release of the oligosaccharides. Chemical methods of releasing *N*-linked oligosaccharides by hydrazinolysis[27] or *O*-linked oligosaccharides using reductive β -elimination[28] are considered the most generic techniques and can be automated for high throughput glycomic analysis. For *N*-linked analysis, the enzymatic use of PNGase F provides an alternative to the chemical hydrazinolysis, but with some limitations in its ability to release *N*-linked oligosaccharides that have substitution on the C-3 position of the reducing end GlcNAc. Both hydrazinolysis and PNGase F generates reducing sugars that can be further derivatized on the reducing end with a chromophore or fluorophore in order to increase the detectability of oligosaccharides during separation. Derivatization of the reducing end has also been done in order to improve chromatographic and electrophoretic behaviour as well as improving MS performance[19]. The reductive β -elimination renders released oligosaccharides alditols that cannot be derivatized at the reducing end. Several promising chemical non reductive β -elimination approaches have been suggested to address this [29-32], while enzymatic approaches so far are limited to only basic *O*-linked oligosaccharides with a core 1 disaccharide Gal β 1-3 GalNAc attached to a serine or threonine. Further enrichment for sulfated oligosaccharides needs to be done after the oligosaccharides have been released since there is usually a mixture of neutral, carboxyl containing (sialic acid and hexuronic acid) and anionic inorganic residues (sulfate and phosphate). It is not uncommon that sulfated oligosaccharides are the least abundant.

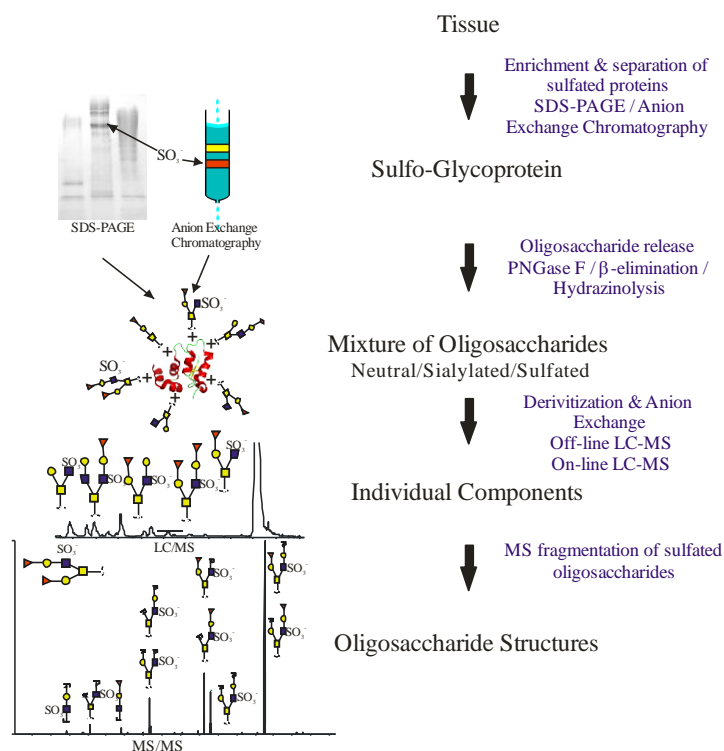


Fig.(2) Generic sulfoglycomic platform for MS analysis of *N*-linked and *O*-linked sulfated oligosaccharides.

This review and perspective of MS based sulfoglycomics will address all the items of the sulfoglycomic platform presented in Figure 2, regarding what is the current state of the art techniques, what areas need to be developed further and what are the common problems and issues that will be difficult to address even with the best sulfoglycomic platforms.

Enrichment of sulfated oligosaccharides

The first step in the scheme for sulfoglycomics platform involves enrichment. The discovery of sulfates on oligosaccharides resembles the history of protein phosphorylation, where the sensitive detection with radioactive labeling is foregoing the actual structural assignment of sulfated oligosaccharides. Sulfated *O*-linked and *N*-linked oligosaccharides were only identified sporadically when detected amongst unsulfated structures in early research, but it was not until enrichment was utilized on a more regular basis that sulfated oligosaccharides could consistently be detected and analyzed with MS. Using anion exchange for fractionation of complex oligosaccharides enabled subsequent purification of sulfated oligosaccharides from various sources using ^1H NMR[33, 34]. This early work in the 80's and 90's serves as inspiration to how current analysis of sulfated oligosaccharides is performed. Anion exchange chromatography has now been adopted as one of the main pathways for enrichment of sulfated oligosaccharides for analysis. With both sulfated and carboxylic acid containing (e.g. sialylated) oligosaccharides retarded on the column, this problem can be resolved using sialidase treatment prior to applying the samples to anion exchange chromatography (Fig. 3). This approach can be utilized in order to detect sulfated structures that without enrichment would be in low abundance amongst a mixture containing abundant neutral and sialylated structures. With ion suppression in the MALDI-MS[35], without enrichment sulfated structures may not even be detected. In the case of LC-MS, the ion suppression may be less of a problem, but the low ion intensity compared to other structures will prevent it from being automatically selected for sequencing by LC-MSⁿ. Multiple LC-MS runs will need to be performed to target these structures separately, where valuable biological samples can be wasted.

Table 2. MS of sulfated Oligosaccharide Conjugates

Type	Derivatisation	Enrichment	Chromatography		MS Mode	Inionization Source	MS	MS ⁿ	Ref
<i>N,O</i> -linked	Free, Alditols	Anion exchange			+/-	ESI	FT-ICR	CID	[36]
<i>O</i> -linked	Free, Alditols	Gel Electrophoresis & Anion Exchange	Online	PGC	-	ESI	IT	CID	[37]
<i>O</i> -linked	Free, Alditols		Nano-Online	PGC	-	ESI	IT	CID	[38]
<i>O</i> -linked	Free, Alditols	Anion Exchange	Capillary & Nano Online	PGC	-	ESI	IT	CID	[37]
<i>O</i> -linked	Free, Alditols		Online	PGC/Amino	-	ESI	IT	CID	[39]
<i>O</i> -linked	Free, Alditols		Online	PGC	-	ESI	IT	CID	[40]
<i>O</i> -linked	AB labelling	Anion Exchange	Offline	Anion Exchange	+/-	MALDI	TOF		[41]
<i>O</i> -linked	Synthetic glycopeptides		Offline	Size Exclusion	-	MALDI	TOF		[42]
<i>O</i> -linked	Free, Alditols	Size Exclusion	Online	Anion Exchange	+	FAB	2 Sector		[34]
<i>O</i> -linked	Free, Alditols		Online	Amino	-	FAB	2 Sector		[33]
<i>O</i> -linked	Per-deuteracetylated	Derivatization/ Anion Exchange	Online		-	FAB	4 Sector	CID	[40]
<i>O</i> -linked	Peracetylated	Derivatization/ Anion Exchange	Online		-	FAB	4 Sector	CID	[21]
<i>N</i> -linked	Free, Alditols	PGC	Online	PGC	+/-	ESI	Q-TOF	CID	[26]
<i>N</i> -linked	Free, Alditols	PGC	Online	PGC	+/-	ESI	Q-TOF	CID	[43]
<i>N</i> -linked	De-Sialylated		Online	PGC	+	ESI	QQQ	CID	[44]
<i>N</i> -linked	PA	Anion Exchange	Offline	HILIC	+/-	MALDI	TOF	CID	[45]
<i>N</i> -linked	De-Sialylated		Direct		-	MALDI	TOF	CID	[46]
<i>N</i> -linked	PA Label	Size Exclusion	Offline	Cation Exchange	+	MALDI	TOF		[47]
<i>N</i> -linked	Permethylation	Reverse phase	Direct		+	MALDI	TOF	CID	[48]
<i>N</i> -linked	Permethylation	Anion Exchange	Direct		+/-	FAB			[49]
<i>N</i> -linked (nonsulfated)	AB	HILIC	Online	HILIC	+	ESI	TOF	CID	[50]
Glycopeptide	Ion Pairing		Direct		+	ESI	IT	CID	[51]
Glycopeptide	Ion Pairing		Direct		+	ESI	FTICR	CID	[52]
Glycopeptide	Ion Pairing		Direct		+	ESI	Q-TOF	CID	[53]
Glycopeptide	"Sulfate Emerging"	Anion exchange	Offline	Reverse phase HPLC	-	MALDI	TOF	-	[54]
Glycoprotein	PA Label	Size Exclusion	Offline	Anion Exchange	+	MALDI	TOF	CID	[55]
Standards	Free, Alditols		Direct		-	ESI	IT	CID	[56]
Standards	Free, Alditols	Size Exclusion	Offline	Size Exclusion	-	ESI	FTICR	EDD	[57]
Standards	Free, Alditols	Gel electrophoresis	Online	PGC	-	ESI	IT	CID	[58]
Standards	Free, Alditols		Nano	Amide	-	ESI	Q-TOF	CID	[59]

Table 2 Contd...

Type	Derivatisation	Enrichment	Chromatography		MS Mode	Inonization Source	MS	MS ⁿ	Ref
Standards	Free, Alditols				-	ESI	IT	CID	[60]
Standards	Free, Alditols				-	ESI	FTICR	EDD	[61]
Standards	Free, Alditols				-	ESI	FTICR	EDD	[62]
Standards	Methyl Esterfication	Size Exclusion			-	Nano-ESI	Q-TOF	CID	[63]
Standards	Ion Pairing		Online	Capillary zone Electrophoresis	+	ESI		CID	[64]
Standards	Ion Pairing		Direct		+	ESI	IT	CID	[65]
Standards	Free, Alditols		Direct		+/-	MALDI	TOF	x	[66]
Standards	Free, Alditols		Direct		+/-	MALDI	TOF	CID	[67]
Standards	Free, Alditols	Gel filtration	Offline	Anion Exhcange	-	MALDI	TOF/TOF	CID	[68]
Standards	Free, Alditols	Size exclusion & Gel Filtration	Direct		-	MALDI	TOF	x	[69]

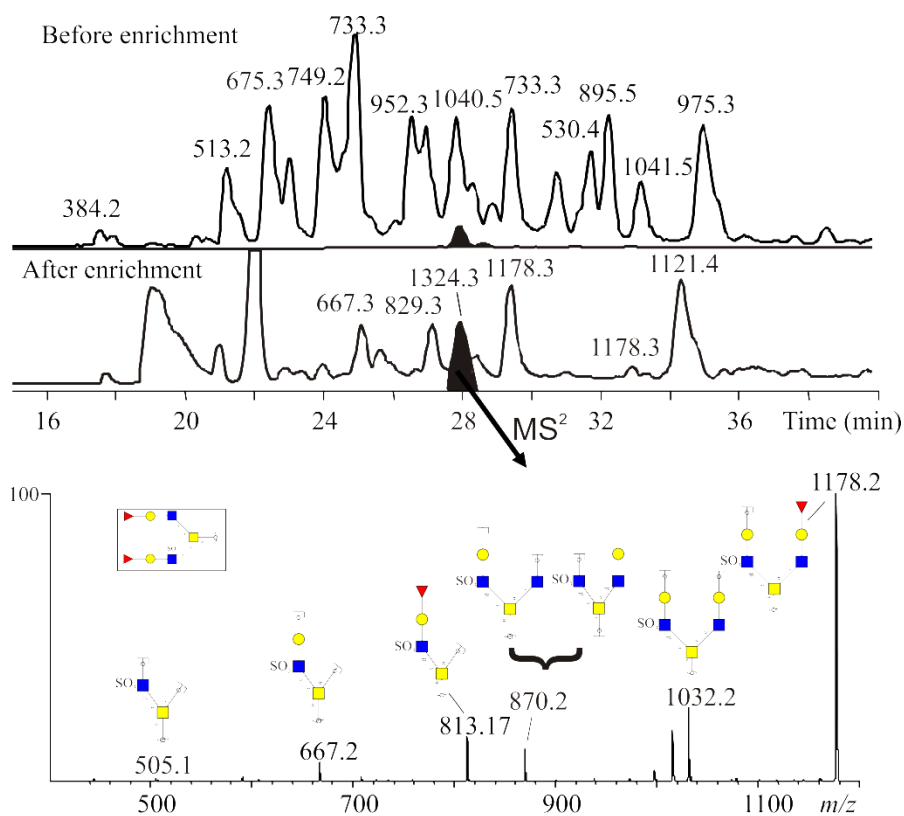


Fig.(3) Base peak chromatogram of salivary O-linked oligosaccharides before and after enrichment for sulfated structures. Base peak chromatograms of total O-linked oligosaccharides with major [M – H]⁻ ions of neutral and sialylated peaks labeled, and sulfated O-linked oligosaccharides after desialylation and anion exchange enrichment. In the chromatograms the [M – H]⁻-component from a monosulfated heptasaccharide is indicated and is successfully characterized by LC-MS². In the lower chromatogram major sulfated [M – H]⁻ ions are labeled. Symbols see Fig 1.

Enrichment of sulfated oligosaccharides without anion exchange

Enrichment of sulfated oligosaccharides with other types of chromatographic material than an anion exchange could also be perceived. An interesting approach was presented using strong cation exchange on a gel permeation media [70]. With the smaller volume for sulfates in this material due to repulsion from the chromatographic media, sulfated structures were found in the excluded fraction, while other oligosaccharides were retarded on the column. This approach was proposed to be effective in separating not only sulfated from neutral oligosaccharides but also from other acidic oligosaccharides containing sialic acid and hexuronic acid [70]. Gel permeation chromatography has been used to isolate pulmonary high molecular mass sulfated oligosaccharides [71]. In proteomics, other types of chromatographic media have been used to enrich for various types of acidic peptides, including phosphates. Methods such as graphitized carbon and HILIC have successfully isolated phosphopeptides and have also been used for separation of sulfated oligosaccharides. The selectivity of the former makes it unlikely that it can be used as a generic method for specific isolation of sulfated oligosaccharides from other oligosaccharides. Further investigation of the use of HILIC will need to be performed to investigate its usefulness for native oligosaccharides, but both HILIC and reversed phase fractionation appears to be useful for permethylated sulfated oligosaccharides [72]. Chromatographic media made from the mineral rutile (TiO_2) are widely used for phosphoproteomics [73], and its affinity for acidic compounds indicates that it may also have a use for enrichment of sulfated oligosaccharides.

Enrichment of sulfated oligosaccharides after derivatization

An elegant method to retain the information of the sulfate and sialic acid residues without the requirement for sialidase treatment prior to enrichment was devised by chemically converting sialic acid into neutral esters. This was followed by anion exchange chromatography of oligosaccharides. With the biological importance of oligosaccharide epitopes containing both sulfate and sialic acid, this type of approach may turn out to be a fruitful way to look at the sulfoglycome. The ester functionality is also easy to remove, allowing the analytical and biological questions of isolated sulfo sialo oligosaccharides or mixtures of oligosaccharides to be addressed. With esterification performed on-column, the separation of neutral, sialylated and sulfated structures could be performed in a one-step procedure [74]. Similar to this approach is the technique devised for enrichment of sulfated glycopeptides [54]. The approach, dubbed a "sulfate emerging" procedure, involves controlling the net charge of sulfated glycopeptides allowing sulfated glycans to be purified and enriched by ion exchange chromatography. By accentuating the negative charge of the sulfate group as well as quenching the negative charges from non-sulfated charged moieties that reside on the glycopeptides, such as carboxyl of sialic acid residues or the C-terminal of the peptide, the oligosaccharide receives its charge exclusively from the sulfated ion. This was achieved by chemical modifications of tryptic glycopeptides, where the C-terminal lysine and arginine residues were removed by carboxypeptidase B. The glycopeptides were desialylated with sialidase A and the carboxyl groups that reside on the peptides were modified with acetohydrazide to remove the negative charges. The sulfated modified glycopeptides, together with peptides containing other anionic organic modifier (e. g. sulfopeptides and phosphopeptides) can easily be isolated using anion exchange chromatography. The enriched glycopeptides can thus be analysed by MS without interference from other charged glycopeptides. This method has been used to enrich and analyse sulfated glycopeptides from complex mixtures where the sulfated structures were claimed to be in low abundance.

With the development of improved methods for permethylation, as described in a later paragraph, the conversion of sialic acid residues into neutral permethylated esters will occur. Hence, in combination with anion exchange this approach has a potential to become one of the main analytical platforms for sulfoglycomics [72]. Since the permethylation also improves MS^n fragmentation interpretation this approach is suitable for de-novo interpretation of novel sulfated oligosaccharide epitopes. When combined with chemical desulfation, it has been suggested that could be a good pathway to address sulfation analysis as the "hole" left after desulfation can serve as a marker for the sulfate position. This "hole" can be deuteromethyl labeled [75] and analyzed in positive ion mode [48].

MS detection of released sulfated oligosaccharides.

A generic method to ionize and detect isolated sulfated oligosaccharides is an important key to the success of the sulfoglycomic research. High quality MS-spectrum of mixtures of sulfated oligosaccharides provides information about composition, sulfate level and provides hints about branching and extension. Ionization of sulfated oligosaccharides is also the prerequisite to be able to perform fragmentation analysis for structural determination. Sulfated oligosaccharides can be analyzed both in positive and negative ion mode and both ESI and MALDI have been utilized (Table 2).

Electrospray is the preferential ionization method for the analysis of sulfated oligosaccharides since it has shown to be less susceptible to concomitant dissociation of sulfate from the oligosaccharide chain. ESI, with or without LC, has been used for the analysis of both *N*-linked and *O*-linked sulfated oligosaccharides (Table 2). A comparison of the influence of organic modifiers in ESI shows that the sensitivity of neutral, sialylated and sulfated oligosaccharides were greatly influenced by the concentration of organic modifier in negative ion mode, whereas it was less dependent in the positive ion mode[43]. Response in positive ion mode was also shown to be reasonably independent of the amount of sialylation, but sulfation significantly suppressed the signal intensity[43]. Although negative ion mode ESI-MS has been used extensively with capillary flow, nanoflow MS has also been explored as a possible platform for the analysis of sulfated oligosaccharides [37, 38]. It offers advantages over its traditional capillary flow counterpart, most notably reduced sample consumption due to increased sensitivity. Good quality MS and MSⁿ of sulfated oligosaccharides is obtainable from 1-3 fmol of analyte in negative ion mode. The observed tenfold increase in sensitivity compared to capillary MS provides good structural elucidation of mono, di- and tri- sulfated glycans isolated from mucins. Interestingly it was shown that increased pH in the nanoflow provided increased sensitivity for sulfated oligosaccharides compared to neutral ones[38].

The traditional MALDI matrices available for biomolecules have typically been selected for their performance in positive ion mode. With the negative charge introduced on sulfated oligosaccharides, MALDI matrices for negative ion mode have been developed (Table 3). However no consensus optimized molecular feature for negative ion mode matrices has been reached. The best current matrices in negative ion mode appear to be the ones also used in positive ion mode, doped or mixed with chemical modifiers. Negative ion mode has shown to be devoid of some of the problems associated with positive ion mode MALDI. The current challenges in negative ion mode MALDI of sulfated oligosaccharides are instead low sensitivity and signal suppression from the less acidic species[35].

Comparing Electrospray with MALDI, a fundamental MS difference between MALDI and ESI is that the former is dominated by singly charged pseudomolecular ions while the latter generates multiply charged ions. Positive ion mode MALDI of oligosaccharides with several acidic groups necessitates the formation of cationic metal adducts to ensure there is a net positive charge. This can make the interpretation of the MS spectra more difficult, where for instance the difference between potassium and sodium adducts differ by 16 amu, which is the same as the difference between deoxyhexose (Fucose) and hexose (Mannose, Glucose, Galactose) as well as *N*-Acetyl and *N*-Glycolyl neuraminic acid. For singly and doubly sulfated oligosaccharides, this appears to be less of a problem (Fig. 4). The composition of sulfated oligosaccharides can be assigned both in positive and negative ion mode MALDI, where oligosaccharides can be detected as singly charged pseudomolecular ions either as $[M-H+2Na]^+$ and $[M-2H+3Na]^+$ or $[M-H]^+$, $[M-2H+Na]^+$ [26]. In negative ion mode MALDI, the acidic (sulfated) oligosaccharides can selectively be detected (Fig. 4), while in positive ion mode both acidic and neutral oligosaccharide components are detected in a mixture. For glycomics in general (not only for sulfoglycomics and negative ion mode) the most popular matrix, 2,5-dihydroxybenzoic acid (DHB), is known to form inhomogeneous needle shaped crystals [76], an attribute that can be responsible for heterologous ionisation properties leading to poor reproducibility of glycomic analysis. With d-arabinosazone included in DHB as a MALDI matrix, it was shown that the post source decay (PSD) was negligible for loss of sulfate negative mode, and it was possible to locate the position of the sulfate on the non-reducing terminal *N*-acetylgalactosamine residue. Furthermore, the use of an osazone as the matrix was found to be particularly good for anionic glycans as it gave a strong signal with low background[77].

Table 3 Example of MALDI-Matrices Used for Analysis of Sulfated Oligosaccharides.

Mode	Matrix	Conc	Solvent	Ref
+/-	DHB	10mg/ml	50% AcN	[78]
+/-	1,1,3,3-tetramethylguanidine salt of CHCA	~90mg/ml	MeOH	[78]
+/-	1,1,3,3-tetramethylguanidine salt of ρ -coumaric acid	~90mg/ml	MeOH	[78]
+/-	Pyrenemethylguanidine HCl	10mM	50%MeOH	[66]
+/-	6-Aza-2-thiothymine	5 mg/ml	H ₂ O	[79]
-	2-(4-hydroxyphenylazo) benzoic acid /1,1,3,3-tetramethylguanidine	70-90mg/ml	MeOH	[69]
-	2- (4 - hydroxyphenylazo) benzoic acid/Spermine	70-90mg/ml	MeOH	[69]
-	DHB	10mg/ml	Ammonium Citrate	[41]
-	D-arabinosazone	3mg/ml	25% EtOH	[26]
-	THAP/ Norharmine	10mg/ml	50% AcN	[68]
-	1-methylimidazolium-CHCA	70mg/ml	MeOH	[68]
-	CHCA/1,1,3,3-tetramethylguanidine	70-90 mg/ml	MeOH	[67]
+	DHB	10mg/ml	MeOH/CHCl ₃	[26]
+	2,5-DHB/3,4-diaminobenzophenone	10mg/ml	50% ACN/ 0.1% TFA	[80]
+	DHB/Sodium Citrate	10mg/ml	MeOH/Water	[48]
+	DHB/Sodium Citrate	10mg/ml	EtOH/Water	[30]
+	DHB	10mg/ml	50% AcN /0.1% TFA	[41]

Ionic Liquid Matrices for MALDI of sulfated oligosaccharides

With the advent of Ionic liquid Matrices (ILM's), there is an expectation that the dissociation of labile groups experienced with MALDI applications for glycomics will slowly but surely cease, and that the particular sulfate dissociation problem associated with MALDI will not be a factor to consider for the analysis of sulfated oligosaccharides. ILM's consist of a conventional acidic MALDI matrix, such as DHB and an organic base typically aliphatic or aromatic amines [36], e.g. tributylamine, pyridine or 1-methylimidazole. With the inclusion of an organic base, ILM's maintain a relative state of "liquidity" under vacuum conditions[67]. This liquidity improves the homogeneity of the "co-crystallized" matrix-analyte complex, thus improving shot to shot reproducibility of the mass spectra compared to using traditional solid matrices. The trade-off for this increased reproducibility has been that the relative sensitivity of ILM's is reduced compared to traditional matrices [76, 81], with the sensitivity of oligosaccharides restricted to the picomolar level.

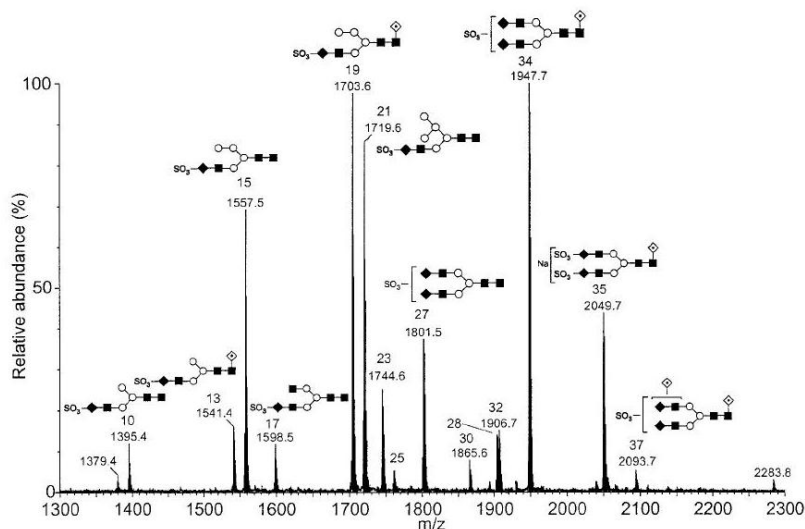


Fig (4) Example of negative ion mode MALDI-MS of sulfated *N*-linked oligosaccharides. MALDI TOF mass spectrum of the *N*-linked glycans from bovine thyroid-stimulating hormone (D-arabinosazone matrix). Key to symbols: □, GlcNAc; ◆, GalNAc; ○, mannose; ◇, fucose. Reprinted from Wheeler *et al.*[26]. Copyright 2001 Academic Press.

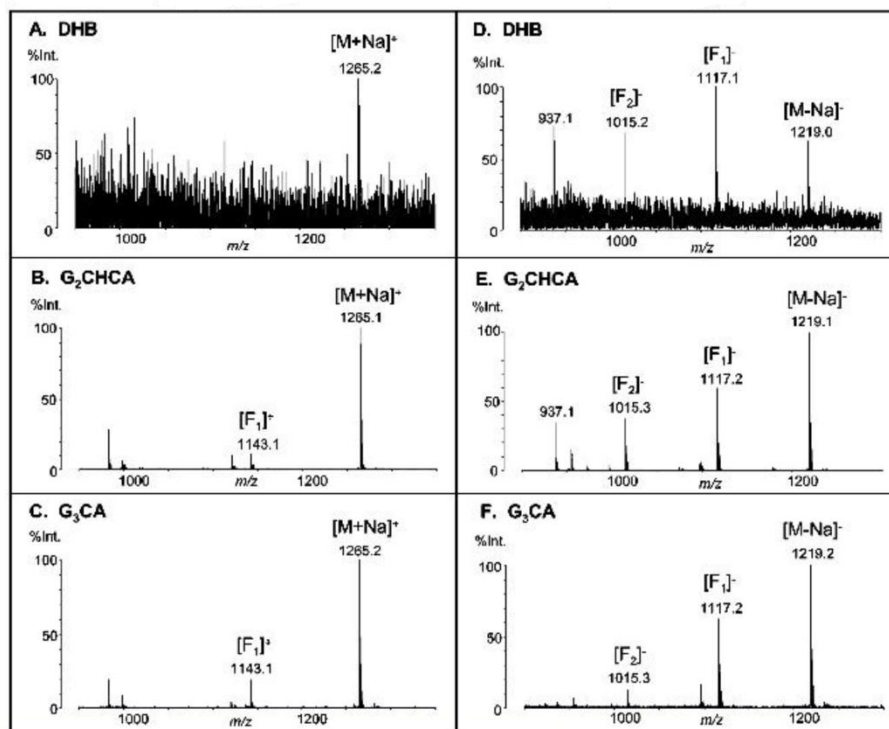
For sulfated oligosaccharides, 1,1,3,3-tetramethylguanidinium (TMG) salts of α -Cyano-4-hydroxycinnamic acid was reported to be particularly good at suppressing the dissociation of the sulfate group[81]. However, the quality of data in negative ion mode was poor compared to positive ion mode and even in positive ion mode the detection limit of > 10 pmol was not impressive. Fukuyama *et al.*[67] used a similar MALDI matrix, though chose a synthesized TMG salt of *p*-coumaric acid (G3CA) as an ILM (Fig. 5). They found G3CA was particularly good at suppressing dissociation of sulfates and sialic acids, and sensitivity was improved with sulfated glycans detected as low as 1 fmol in positive ion mode. They also reported that G3CA was good for preferential detection of sulfated glycopeptides, especially in negative ion mode. An additional indication that ILM's would be good for MALDI of *O*-linked and *N*-linked oligosaccharides was shown using a combination of a crystalline 2-(4-hydroxyphenylazo)benzoic acid (HABA) with TMG. This ILM was particularly good for the detection of sulfated structures in negative ion mode as shown by the analysis of partially digested glycosaminoglycans[69]. These results show potential for the use of MALDI-MS in the investigation of acidic oligosaccharides, especially where high throughput is required.

Isomeric and isobaric challenges for sulfoglycomics

One of the challenges in MS based glycomics is that isomeric structures are frequent. One way to address this is based on complicated identification schemes using multiple MSⁿ fragmentation. Alternatively, oligosaccharides have to be isolated either off-line or on-line prior to MS analysis. With the exponential growth of possible isomers by the increasing number of possible monosaccharide residues or substitutions in glycomic solution space, what was already a difficult task without sulfate substitution, could become an insolvable chromatographic or electrophoretic equation for all potential sulfate positions in a glycomic analysis. Fortunately, for mammalian *N*-linked and *O*-linked oligosaccharides, sulfation appears to be dominated to C-3 on Gal residues, C-4 on GalNAc residues and C-6 on GlcNAc residues, thus limiting the isomeric possibilities and simplifying the interpretation of the mass spectra.

In addition to isomeric challenges, there is also a potential for difficulties in differentiating isobaric peaks. The accuracy of most mass spectrometers used in life science is usually not high enough to allow discrimination between sulfated and phosphorylated species (mass difference of 0.0095 amu). This is important since they are associated with significantly different biological functions. Therefore developing efficient and sensitive methods to discriminate between these structures is essential for understanding their roles. The mass to charge differentiation becomes even more challenging since underivatized sulfated and phosphated oligosaccharides are typically multiply charged. Although a high mass accuracy MS such as a FTICR [65] would be able to

discriminate between these small mass differences, confident assignment can also be achieved by their different fragmentation. One such method developed by Zhang *et al.*[51] utilized the different proton affinities of sulfate and phosphates and therefore their different chemical reactivities with an ion-pairing reagent. By interpretation of the fragment spectra of sulfated and phosphorylated standards obtained from low resolution mass spectrometers, it was possible to differentiate between the two species. They also noted an enhanced mass spectral signal from improved ionisation properties afforded by the use of ion-pairing reagents and found their method was particularly useful for detecting low abundant sulfated oligosaccharides groups within a complex mixture



Fig(5) Ion Liquid Matrices (ILMs) for MALDI-MS for highly sulfated oligosaccharides. Trisulfated neocarrahexaose oligosaccharide detected in positive and negative ion mode using conventional DHB (A and D) and the ILMs ditetramethylguanidine α -cyano-4-hydroxycinnamate (B and E), and tritetramethylguanidine

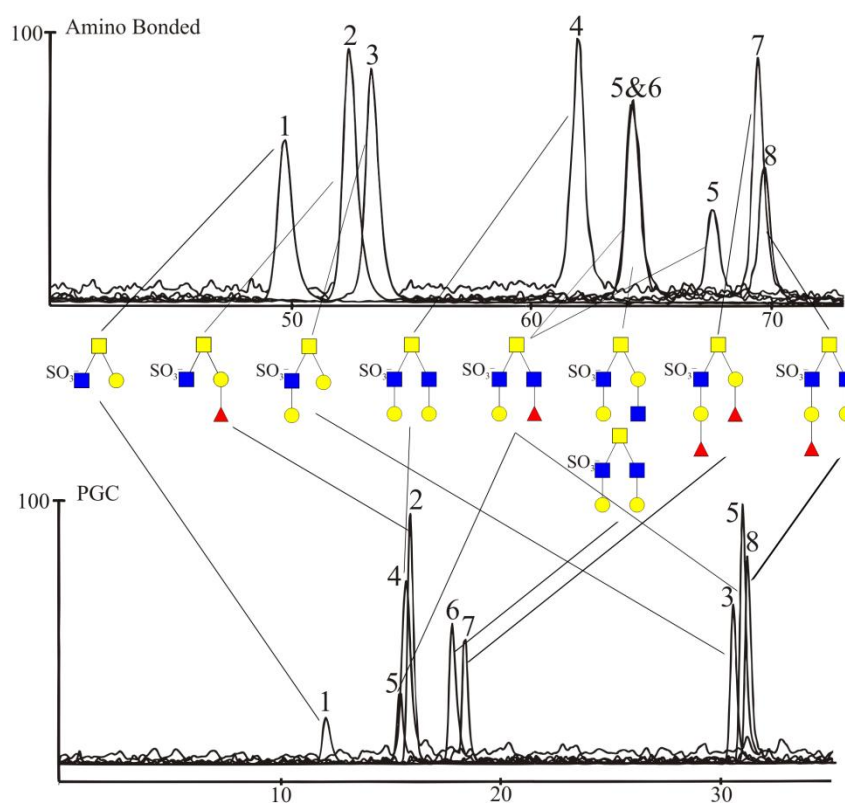
Peracetylation and permethylation

Permethylation and peracetylation are techniques that originally were developed in order to make monosaccharides and oligosaccharides more volatile and allow them to be analyzed by GC and GC-MS without decomposing. Peracetylation is the addition of an acetyl group onto the hydroxyl residues of a carbohydrate[82] and is best carried out by base-catalysis with either pyridine or methylimidazole, hence it is suitable for labile sulfated oligosaccharides[21]. Reduction of monosaccharides and oligosaccharides is usually carried out to deal with the anomeric conversion at the reducing end during peracetylation[83]. The problem with MS analysis of peracetylated oligosaccharides is that a hexose and an *N*-acetylhexosamine will only differ by 1 amu after derivatization. An alternative use of perdeutero-acetylation increases this difference to 4 amu and has successfully been used to identify *O*-linked sulfated oligosaccharides [20, 22]. For permethylation of sulfated oligosaccharides, the protocol developed in the 1960's was based on a two-step procedure involving an initial generation of a DMSO carbanion that could extract protons from the oligosaccharide hydroxyls, followed by a nucleophilic attack on all the hydroxyls with methyl iodide[84]. This procedure was quite involved and the introduction of a one-pot reaction using solid sodium/potassium hydroxide and methyl iodide made permethylation a standard technique in many glycoanalytical laboratories[85]. Unfortunately while the early carbanion method could successfully be applied to sulfated oligosaccharides [49], the novel more popular one pot reaction could not. It was realized that it was not the lability of the sulfate group, but the actual work up procedure after derivatization that was problematic [48, 80]. Exchanging the organic/water partitioning work up

procedure with reversed phase clean up meant that sulfated oligosaccharides could be recovered and analyzed. This approach is now utilized for analyzing both *N*-linked and *O*-linked sulfated oligosaccharides. Further refinement of the platforms for permethylated sulfated oligosaccharides analysis in positive ion mode after chemical desulfation has been suggested[48].

LC-MS for analysis of sulfated oligosaccharides

After enrichment of the sulfoglycome and detection from the mixture of sulfated oligosaccharides, sulfated structures will have to be identified. An integrated LC-MS system has the potential to be a very effective tool for the analysis of sulfated oligosaccharides as it can resolve some of the isomeric issues. Oligosaccharides released with their reducing-end intact exist in at least two dominating forms due to the anomeric configuration. To avoid anomeric separation, sulfated oligosaccharides are preferably analysed by LC-MS after reduction of the oligosaccharides to alditols or after reductive amination[19]. Independent of running in normal or reverse phase style separation, the eluting solvents must be compatible with the mass spectrometer. The selectivity of the chromatography will of course differ depending on hydrogen bonding interaction (e.g. normal phase amine or amide columns) or hydrophobic interaction (PGC). This is illustrated in Figure 6 [39]. From this figure one can see that the main selectivity of normal phase is the size of the oligosaccharides, while graphitized carbon selectivity is more complex based on hydrophobic region of the molecule. Typically structures containing type 2 *N*-acetylglucosamine and blood group H are generally more retarded, and Lewis type structures and type 1 *N*-acetylglucosamine containing structures are less retarded on a graphitized carbon column. With complex mixtures of sulfated oligosaccharides a dual chromatographic setup of graphitized carbon and normal phase would be required to resolve all isomers[39].



Fig(6) LC-MS of *O*-linked sulfated oligosaccharides. Extracted ion chromatograms of sulfated *O*-linked porcine gastric oligosaccharide alditols, separated on an (a) amino-bonded column and (b) Porous graphitized carbon column, and detected as (M-H)⁻ ions with single quadrupole ESI-MS. Modified from Thomsson *et al.*[39]. Copyright 1999 Elsevier B.V.

Different normal phases have successfully been utilized for analysing oligosaccharides and have the potential for successful integration with ESI-MS [86, 87]. Since the mass spectrometer is already separating on size, the extra

dimension provided by chromatography should target isomeric separation. The LC-concentration effect and the use of chromatography for removing interfering salts and other contaminating compounds means that LC-MS is more sensitive compared to static ESI-MS. LC-MS is also easier to automate, and LC-MSⁿ can be done “on the fly” for sulfated oligosaccharides to generate structural information[37]. Stationary phases of silica with amine, aminopropyl groups or underivatized silica in the case of weak ionic columns as well as non-ionic amide or diol columns would all provide isomeric separation of oligosaccharides and be ideal for LC-MS setup [86, 87]. One of the first examples of connecting LC with MS for successful accumulation of MS and MSⁿ for structural characterisation of *O*-linked sulfated oligosaccharides was described using amino-bonded normal phase chromatography connected to a QTOF mass spectrometer[88]. Reverse phase type chromatography (graphitized carbon) has been used extensively for LC-MS analysis of sulfated oligosaccharides as it can separate both neutral and acidic oligosaccharides based on size as well as providing isomeric separation[14, 15, 24, 25, 37, 38, 43, 44, 89, 90]. When PGC LC is used in conjunction with a slightly alkaline elution buffer (such as ammonium bicarbonate) it is an effective chromatography method for the resolution of sulfated oligosaccharides [25, 37, 38, 44, 90]. When using acidic buffers (ammonium formate, pH 3.0) neutral glycans are unaffected but sialylated and sulfated species are retained somewhat stronger[43]. Positive ion mode has been shown to significantly suppress the signal of native sulfated oligosaccharides compared to neutral and sialylated structures in LC-MS using graphitized carbon[43].

With the glycomic MS moving into miniaturization, the nano LC-MS development for oligosaccharide analysis still needs significant development in order to be applicable to sulfoglycomics. Capillary LC-MS, using both normal and PGC stationary phases, is becoming routine with sub-picomole detection of oligosaccharides, including sulfated ones. Both positive and negative ion mode LC-MS are utilized in glycomics, although the preference of negative ion mode ESI-MS for sulfated oligosaccharides puts constraints on nano LC-MS. In high water conditions the reduction of water into generating hydrogen gas could become significant in negative ion nano flow. The standard potential for reduction of water into hydrogen is 0 V in acidic conditions and -0.83 V in alkaline, while the oxidation into oxygen is +0.40 V (alkaline) and +1.23 V (acidic) indicating that positive ion mode will have fewer problems with the water redox potential in acidic conditions, while negative mode is benefited by the high alkaline conditions. Since normal phase chromatography has less water in the mobile phase it can successfully be adapted to negative ion LC-MS despite the redox constraints. A method developed using an amide-silica HILIC on a nano chip based format connected to an ESI-QTOF-MS was demonstrated to be particularly useful for the analysis of chondroitin sulfate derived oligosaccharides [59, 91]. An alternative approach to counteract the reduction of water in the high water content used in PGC nano LC-MS, increased pH using ammonium bicarbonate or ammonium hydroxide has been used successfully for analysis of *O*-linked sulfated oligosaccharides[37, 38].

Off-line separation of sulfated oligosaccharides is commonly performed, but has rarely been utilized to address high throughput sulfoglycomics, but rather low throughput thorough characterization using MS in combination with other analytical techniques. However, the potential of being utilized for sulfoglycomics has been demonstrated using off-line PGC and MALDI-MS. Lebrilla *et al.* [92] separated *O*-linked oligosaccharides with stepwise elution of 10%-20% acetonitrile. The subsequent fractions were analyzed by MALDI-FT-ICR-MS using a matrix of DHB in positive ion mode.

Fragmentation analysis of sulfated oligosaccharides

The sulfate groups on oligosaccharides provide negative charges to the compounds; therefore negative ion mode fragmentation tends to be the method of choice. Despite that, much of the original fragmentation analyses of sulfated oligosaccharides were performed in positive mode using FAB. In positive mode the sulfated oligosaccharides are usually analyzed as their sodiated adducts, with at least two sodium ions available on the parent ion (one to counteract the negatively charged sulfate and one to generate the charge of the parent ion). The sulfated oligosaccharides provide fragment ions similar to neutral oligosaccharides, where the localization of the charge-carrying sodium dictates the fragmentation. In negative ion mode, the CID displays classical charge remote fragmentation[93]; this is independent of whether it is high or low energy collision. High energy

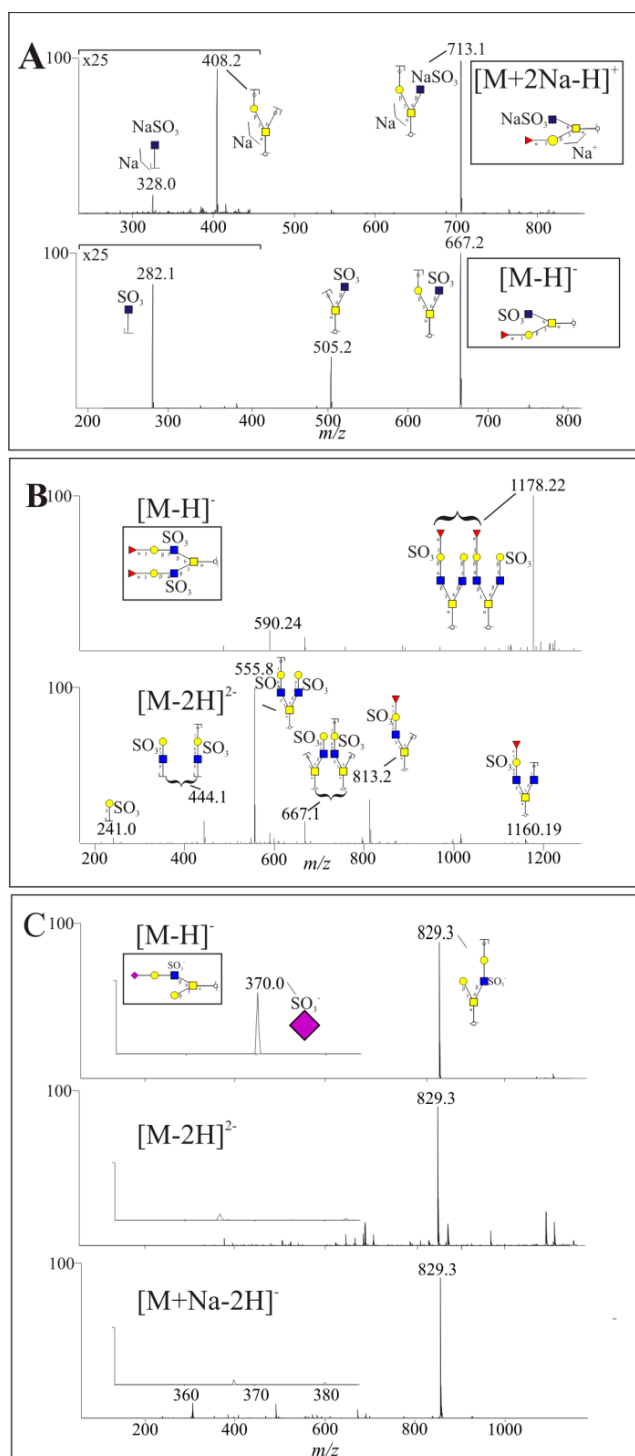


Fig.(7): CID fragmentation of sulfated oligosaccharides. Mass Spectra of a sulfated *O*-linked tetrasaccharide negative ion mode as $[M-H]^-$ and positive ion mode as $[M+2Na-H]^+$ (a). Fragmentation of a doubly sulfated *O*-linked hexasaccharide showing improved quality of the fragmentation of the doubly charged $[M-2H]^{2-}$ -ion compared to the $[M-H]^-$ -ion (b). CID induced migration of sulfate (m/z 370) in a sulfated sialylated *O*-linked pentasaccharide detected fragmentation of the $[M-H]^-$ -ion but not the $[M-2H]^{2-}$ -ion or the $[M+Na-2H]^-$ -ion (c). Symbols see Fig 1.

collision has the tendency to provide more cross ring fragmentation. With sulfated *O*-linked oligosaccharide alditols, the main type of fragmentation is the $^{1,5}X$ cross ring charge remote fragments both for native and peracetylated oligosaccharides[21]. Low energy collision of both *N*-linked and *O*-linked oligosaccharides produces mostly reducing end Y type fragments containing the sulfate group or non-reducing end B-type ions containing the sulfate group, also by charge remote type fragmentation[37, 94]. In Figure 7a CID fragmentation of a sulfated *O*-linked tetrasaccharide is shown and it can be seen that positive ion mode fragmentation is directed towards the reducing end, indicating that the charge-carrying sodium is coordinated mainly at the reducing end residues, whereas in the negative ion mode fragmentation spectra it is instead directed towards the negatively charged sulfated GlcNAc residue on the C-6 branch of the reducing end GalNAcol illustrating the point of charge remote fragmentation in negative ion mode.

Both high and low energy CID provides spectra of native oligosaccharides that have been shown to be able to determine the sequence of sulfated oligosaccharides and identification of sulfated residue and potentially the sulfation position [20-22, 37, 94]. The level of sulfation of *N*- and *O*-linked oligosaccharides is usually quite low and for singly sulfated, singly charged oligosaccharides this approach is successful. For oligosaccharides with multiple acidic residues in the oligosaccharides such as sulfate and sialic acid, or multiple sulfates it appears that multiply charged or sodiated singly charged oligosaccharides provide more informative fragmentation than their singly charged parent ion counterpart (Fig. 7b). The latter are dominated by fragment ions from the loss of acidic residues. In the case of sialylated sulfated oligosaccharides, esterification of the sialic acid on less negatively charged parent ions are likely to provide more information from the fragmentation spectra [95].

There is the possibility to generate more information by multiple stage fragmentation MS^n of sulfated oligosaccharides[37]. Alternatively other types of fragmentation, derivatization and ionization modes have been tested in order to increase the amount of structural information obtained from MS^n sulfated oligosaccharides. Comparison between CID fragmentation and IRMPD, showed that the two

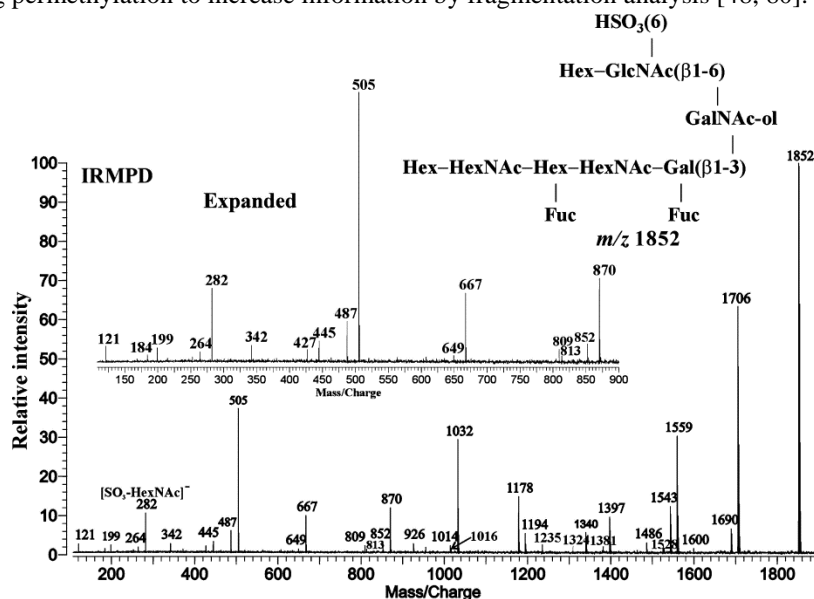
methods are similar for *O*-linked sulfated oligosaccharides (Fig. 8), showing that efficient sequencing of sulfated oligosaccharides could also be achieved using MALDI FT- ICR MS.

Migration of sulfate during fragmentation of sulfated oligosaccharides

A challenging aspect with MS analysis of sulfated glycans concerns the labile nature of the sulfate groups. This problem is obvious in the dissociation of sulfate seen in MS and MSⁿ, and can be detrimental both for appreciating the level of sulfation when analysing oligosaccharide mixtures, and in determining the structure of sulfated oligosaccharides by fragmentation. Even more concerning is the phenomenon observed where the sulfate group is migrating from one monosaccharide residue to a different monosaccharide residue on the oligosaccharides during fragmentation [89]. The analysis of sulfated oligosaccharides by MS in negative ion mode can provide non-conclusive information on the position of the sulfate due to the migration of the sulfate ion (Fig. 7c) [89]. In the figure it can be seen that the sulfate is capable of migrating to the sialic acid residue (*m/z* 370) during the collision of the singly charged [M – H] ion. In order to avoid this problem, the sulfated oligosaccharide are preferentially analysed after CID at precursor ions without acidic proton. The migration was not detected in the fragment spectrum of the doubly charged parent ion or in the fragment spectrum of the parent ion of the singly charged sodium adduct. This phenomenon is similar to that observed in positive ion mode, where the presence of an acidic proton is promoting the migration of fucose after fragmentation of the [M+H]⁺ ion but not of the [M+Na]⁺ ion[96]. Alternatively, permethylating the sulfated oligosaccharide would prevent migration occurring. However as mentioned, the permethylation procedure can be complicated for sulfated oligosaccharides due to the labile nature of the sulfate and such an approach may not suit every laboratory.

Fragmentation of derivatized sulfated oligosaccharides

With regards to derivatization, a methodical investigation of reducing end derivatization to direct fragmentation for sulfated oligosaccharides is still lacking. The use of reducing end derivatization has not been actively used to improve or alter the fragmentation of sulfated oligosaccharides, but rather as part of a general glycomic protocol. Fragmentation of peracetylated oligosaccharides has been shown to generate characteristic sequence ions in negative ion mode, which can be used both to determine the oligosaccharide sequence and sulfate position. However, the labile acetyl group makes the fragmentation inefficient, since abundant fragment ions are generated from the loss of acetyl groups from the parent ion. With the development of novel methods of permethylation, the increased stability of the methyl groups compared to the acetyl groups can now be investigated using permethylation to increase information by fragmentation analysis [48, 80].



Fig(8). IRMPD of a singly charged mono sulfated *O*-linked deca-saccharide. Precursor selected after negative ion MALDI-FTICR-MS. Reproduced from Zhang *et al.*[97]. Copyright 2005 American Chemical Society

It has been indicated that negative ion CID post permethylation may only provide a limited amount of charge remote fragmentation of *N*-linked sulfated oligosaccharides, while positive ion mode using high energy CID MALDI-MS of the sodium adducts provides a substantial amount of information about linkage and branching of both *N*-linked and *O*-linked oligosaccharides ([80] and Fig. 9).

The loss of sulfates that used to be a problem in positive ion FAB-MS of native sulfated oligosaccharides[34] appears not to be a problem in CID after permethylation. This suggests that an approach for sulfoglycomics involving permethylation has the potential of being a reliable and reproducible platform for sulfoglycomics [48, 80]. Chemical desulfation after permethylation, followed by derivatization of the newly generated free hydroxyl with a trideuteromethyl group as a marker for the sulfate position has also been suggested. This approach has the benefit that it can utilize the vast knowledge of fragmentation of permethylated oligosaccharides without the difficulties associated with the sulfate group. MS of permethylated oligosaccharides is currently one of the main approaches for glycomic analysis, and the permethylation/desulfation method would be easy to fit into current glycobioinformatic platforms[48]. The development of a simple permethylation approach for sulfated oligosaccharides will complement current methodologies for analyzing native and reducing end derivatized oligosaccharides with MS.

Alternative fragmentation pathways for sulfated oligosaccharides

Electron transfer dissociation (ETD), electron capture dissociation (ECD) and electron detachment dissociation (EDD) have been investigated as possible alternative fragmentation methods as the fragmentation pathways are totally different to conventional CID. The development of these methods can greatly increase their analytical utility for the characterization of sulfated oligosaccharides. All the aforementioned methods are based on a charge-reduced precursor ion containing a radical site, and this leads to the fragmentation processes (Fig. 10). ECD and ETD are performed in positive ion mode and require multiply charged positive precursor ions which can be difficult to obtain for sulfated oligosaccharides, [62]. This can be overcome by the addition of divalent metals such as Ca^{2+} or cations such as Na^+ . EDD is better suited to the analysis of acidic oligosaccharides as it is exclusively used in negative ion mode.

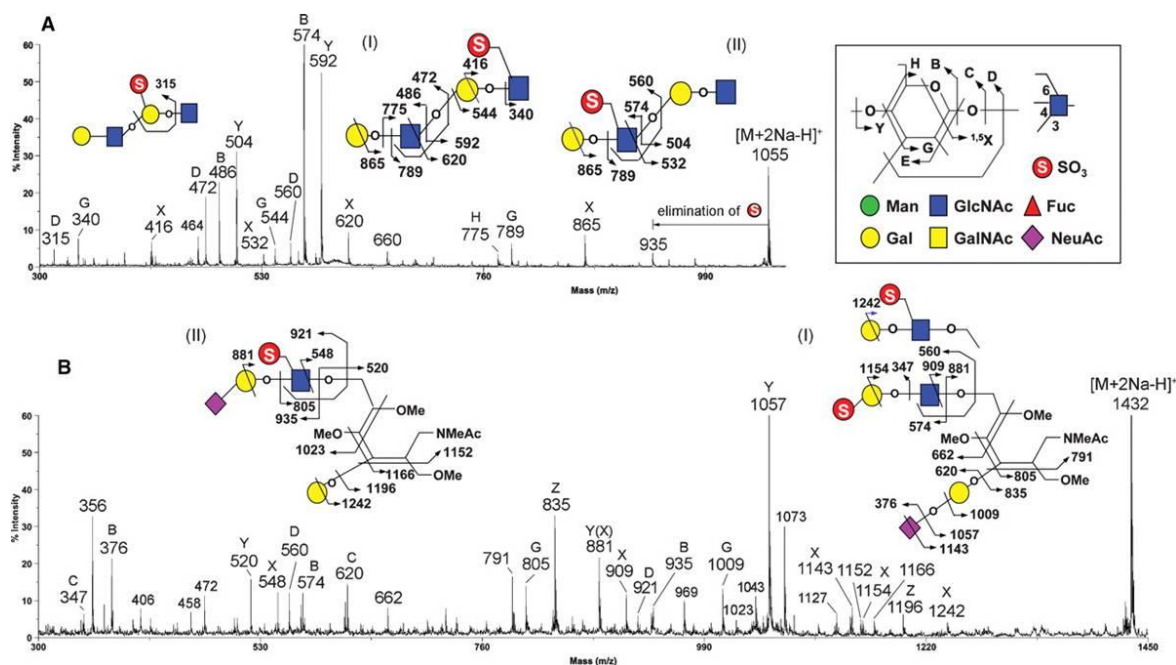


Fig.(9) High energy CID MALDI-MS/MS fragmentation characteristics of permethylated sulfated glycans in the positive-ion mode. Representative high energy MALDI tof-tof CID MS/MS spectra for the permethylated monosulfated diLacNAc derived from partial methanolysates of a tetrasulfated diLacNAc oligosaccharide standard (A) and a sialylated sulfated core 2 *O*-glycan from human ovary cyst mucin (B) are shown here along with schematic drawings of the deduced isomeric structures to illustrate the assignment of the major fragment ions. Reproduced from Yu *et al.*[80]. Copyright 2009 Oxford University Press.

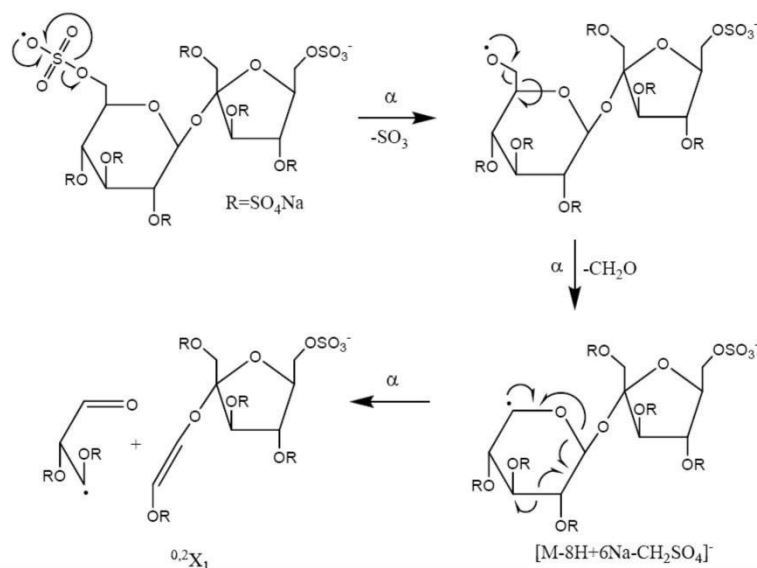


Fig.(10) Schematic diagram showing the fragmentation pathway for EDD. Reproduced from Wolff *et al.* [61]. Copyright 2009 IM Publications LLP.

The future of sulfoglycomics

Sulfation is a ubiquitous modification that we know very little about. Many of the techniques needed to address sulfoglycomics such as sample preparation, sample enrichment, oligosaccharide separation and MS detection have already been developed, but have not been applied to high throughput sulfoglycomics discovery. However in the post genomic era, with improvement in miniaturization and automation that it brings, the life science researcher is now ready to address the area of post translational modifications including sulfation of oligosaccharides. Both MALDI and ESI are effective tools to address the sulfoglycomics questions. The appropriate analytical platforms for sulfoglycomics are now in place to allow for the accumulation of a significant quantity of data. The current bottle neck for sensitive omic analysis in general is neither the access to biological samples nor the generation of MS data, but actually the interpretation of the data. With the increasing quantity of data on sulfated oligosaccharides becoming available, this will be a particular obstacle that will need to be addressed. Software aiding the interpretation of glycomics data may have to be adopted for sulfoglycomics analysis. MS and structural data for sulfoglycomics will need to be stored appropriately and shared in communal databases. Without the symbiotic relationship, as was developed between genomics and proteomics, glycomics and sulfoglycomics will have to rely on experimental data generated by the researchers to build database platforms to aid the development of glycobioinformatics.

List of Abbreviations:

PGC= Porous Graphitized Carbon;

HILIC= Hydrophilic Interaction Liquid Chromatography;

IT=Ion Trap;

QQQ= Triple Quadrupole MS;

LTQ-FTICR= Linear Ion Trap - Fourier Transform Ion Cyclotron Resonance Mass Spectrometer;

CID= Collision Induced Dissociation;

IRMP= Infrared Multiple Photon Dissociation;

PSD= Post Source Decay;

ETD= Electron Transfer Dissociation;

ECD= Electron Capture Dissociation;

EDD= Electron Detachment Dissociation;

TMG= 1,1,3,3-tetramethylguanidium;

CHCA= α -Cyano-4-hydroxycinnamic acid;

TMG= 1,1,3,3-Tetramethylguanidine; G₃CA,

TMG= salt of *p*-Coumaric Acid;

HABA= 2-(4-Hydroxyphenylazo)benzoic acid ;

ILM= Ionic Liquid Matrix,

THAP= 2',4',6'-Trihydroxyacetophenone monohydrate;

DHB= 2,5-Dihydroxybenzoic acid;

PA= Pyridilamine; Ab, Antibody;

Conflict of Interest

The authors declare no conflict of interest. This work was supported by Science Foundation Ireland (06/RFP/CHEF522) and the EU Marie Curie Program (PIRG-GA-2007-205302).

REFERENCES

1. Zaia, J., *On-line separations combined with MS for analysis of glycosaminoglycans*. Mass Spectrometry Reviews, 2009. **28**(2): p. 254-272.
2. Mormann, M., et al., *Analysis of Oversulfation in a Chondroitin Sulfate Oligosaccharide Fraction from Bovine Aorta by Nanoelectrospray Ionization Quadrupole Time-of-Flight and Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry*. Journal of the American Society for Mass Spectrometry, 2007. **18**(2): p. 179-187.
3. Minamisawa, T., K. Suzuki, and J. Hirabayashi, *Multistage mass spectrometric sequencing of keratan sulfate-related oligosaccharides*. Analytical Chemistry, 2006. **78**(3): p. 891-900.
4. Baenziger, J.U. and E.D. Green, *Pituitary glycoprotein hormone oligosaccharides: structure, synthesis and function of the asparagine-linked oligosaccharides on lutropin, follitropin and thyrotropin*. Biochimica et Biophysica Acta, 1988. **947**(2): p. 287-306.
5. Morita, I., et al., *Expression and function of the HNK-1 carbohydrate*. Journal of Biochemistry, 2008. **143**(6): p. 719-724.
6. Brockhausen, I., *Sulphotransferases acting on mucin-type oligosaccharides*. Biochemical Society Transactions, 2003. **31**(2): p. 318-325.
7. Girard, J.P. and T.A. Springer, *High endothelial venules (HEVs): specialized endothelium for lymphocyte migration*. Immunol. Today, 1995. **16**(9): p. 449-457.
8. Uchimura, K. and S.D. Rosen, *Sulfated L-selectin ligands as a therapeutic target in chronic inflammation*. Trends in Immunology, 2006. **27**(12): p. 559-565.
9. Hudson, S.A., et al., *Eosinophil-selective binding and proapoptotic effect in vitro of a synthetic Siglec-8 ligand, polymeric 6'-sulfated sialyl Lewis x*. Journal of Pharmacology and Experimental Therapeutics, 2009. **330**(2): p. 608-612.
10. Roussel, P., G. Lamblin, and P. Degand, *Heterogeneity of the carbohydrate chains of sulfated bronchial glycoproteins isolated from a patient suffering from cystic fibrosis*. J. Biol. Chem., 1975. **250**(6): p. 2114-2122.
11. Boat, T.F. and P.W. Cheng, *Biochemistry of airway mucus secretions*. Federation Proceedings, 1980. **39**(13): p. 3067-3074.
12. Lo-Guidice, J.M., et al., *Sialylation and sulfation of the carbohydrate chains in respiratory mucins from a patient with cystic fibrosis*. Journal of Biological Chemistry, 1994. **269**(29): p. 18794-18813.
13. Davril, M., et al., *The sialylation of bronchial mucins secreted by patients suffering from cystic fibrosis or from chronic bronchitis is related to the severity of airway infection*. Glycobiology, 1999. **9**(3): p. 311-321.
14. Schulz, B.L., et al., *Mucin glycosylation changes in cystic fibrosis lung disease are not manifest in submucosal gland secretions*. Biochemical Journal, 2005. **387**(Pt 3): p. 911-919.
15. Schulz, B.L., et al., *Glycosylation of sputum mucins is altered in cystic fibrosis patients*. Glycobiology, 2007. **17**(7): p. 698-712.
16. Hård, K., et al., *The Asn-linked carbohydrate chains of human Tamm-Horsfall glycoprotein of one male. Novel sulfated and novel N-acetylgalactosamine-containing N-linked carbohydrate chains*. European Journal of Biochemistry, 1992. **209**(3): p. 895-915.
17. Wada, Y., et al., *Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study*. Glycobiology, 2007. **17**(4): p. 411-422.
18. Wada, Y., et al., *Comparison of methods for profiling O-glycosylation: Human Proteome Organisation Human Disease Glycomics/Proteome Initiative multi-institutional study of IgA1*. Molecular and Cellular Proteomics, 2010. **9**(4): p. 719-727.
19. Karlsson, N.G. and N.H. Packer, *Glycomic mass spectrometric analysis and data interpretation tools*, in *Bioinformatics for Glycobiology and Glycomics: An Introduction*, C.-W. von der Lieth, T. Lueteteke, and M. Frank, Editors. 2009, John Wileys & Sons Ltd: Chichester. p. 233-255.
20. Karlsson, N.G., et al., *Molecular characterization of the large heavily glycosylated domain glycopeptide from the rat small intestinal Muc2 mucin*. Glycoconjugate Journal, 1996. **13**(5): p. 823-31.
21. Karlsson, N.G., H. Karlsson, and G.C. Hansson, *Sulphated mucin oligosaccharides from porcine small intestine analysed by four-sector tandem mass spectrometry*. Journal of Mass Spectrometry, 1996. **31**(5): p. 560-572.
22. Karlsson, N.G., et al., *The glycosylation of rat intestinal Muc2 mucin varies between rat strains and the small and large intestine. A study of O-linked oligosaccharides by a mass spectrometric approach*. Journal of Biological Chemistry, 1997. **272**(43): p. 27025-27034.
23. Estrella, R.P., et al., *The glycosylation of human synovial lubricin: implications for its role in inflammation*. Biochemical Journal, 2010: p. in press.

24. Karlsson, N.G. and K.A. Thomsson, *Salivary MUC7 is a major carrier of blood group I type O-linked oligosaccharides serving as the scaffold for sialyl Lewis x*. *Glycobiology*, 2009. **19**(3): p. 288-300.
25. Thomsson, K.A., et al., *MUC5B glycosylation in human saliva reflects blood group and secretor status*. *Glycobiology*, 2005. **15**(8): p. 791-804.
26. Wheeler, S.F. and D.J. Harvey, *Extension of the In-Gel Release Method for Structural Analysis of Neutral and Sialylated N-Linked Glycans to the Analysis of Sulfated Glycans: Application to the Glycans from Bovine Thyroid-Stimulating Hormone*. *Analytical Biochemistry*, 2001. **296**(1): p. 92-100.
27. Wing, D.R., et al., *Use of large-scale hydrazinolysis in the preparation of N-linked oligosaccharide libraries: application to brain tissue*. *Glycoconjugate Journal*, 1992. **9**(6): p. 293-301.
28. Carlson, D.M., *Oligosaccharides isolated from pig submaxillary mucin*. *J.Biol.Chem.*, 1966. **241**: p. 2984-2986.
29. Maniatis, S., H. Zhou, and V. Reinhold, *Rapid de-O-glycosylation concomitant with peptide labeling using microwave radiation and an alkyl amine base*. *Analytical Chemistry*, 2010. **82**(6): p. 2421-2425.
30. Huang, Y., Y. Mechref, and M.V. Novotny, *Microscale nonreductive release of O-linked glycans for subsequent analysis through MALDI mass spectrometry and capillary electrophoresis*. *Analytical Chemistry*, 2001. **73**(24): p. 6063-6069.
31. Karlsson, N.G. and N.H. Packer, *Analysis of O-linked reducing oligosaccharides released by an in-line flow system*. *Analytical Biochemistry*, 2002. **305**(2): p. 173-185.
32. Yamada, K., et al., *Rapid and sensitive analysis of mucin-type glycans using an in-line flow glycan-releasing apparatus*. *Analytical Biochemistry*, 2007. **371**(1): p. 52-61.
33. Strecker, G., et al., *Determination of the structure of sulfated tetrasaccharides and pentasaccharides obtained by alkaline borohydride degradation of hen ovomucin- A fast atom bombardment-mass spectrometric and H-1-NMR spectroscopic study*. *Glycoconjugate Journal*, 1987. **4**(4): p. 329-337.
34. Mawhinney, T.P. and D.L. Chance, *Structural Elucidation by Fast Atom Bombardment Mass Spectrometry of Multisulfated Oligosaccharides Isolated from Human Respiratory Mucous Glycoproteins*. *Journal of Carbohydrate Chemistry*, 1994. **13**(6): p. 825 - 840.
35. An, H.J. and C.B. Lebrilla, *Suppression of sialylated by sulfated oligosaccharides in negative MALDI-FTMS*. *Israel Journal of Chemistry*, 2001. **41**(2): p. 117-128.
36. Petzold, C.J., M.D. Leavell, and J.A. Leary, *Screening and identification of acidic carbohydrates in bovine colostrum by using ion/molecule reactions and Fourier transform ion cyclotron resonance mass spectrometry: Specificity toward phosphorylated complexes*. *Analytical Chemistry*, 2004. **76**(1): p. 203-210.
37. Karlsson, H., et al., *High-throughput and high-sensitivity nano-LC/MS and MS/MS for O-glycan profiling*. *Methods in Molecular Biology*, 2009. **534**: p. 117-131.
38. Thomsson, K.A., et al., *Enhanced detection of sialylated and sulfated glycans with negative ion mode nanoliquid chromatography/mass spectrometry at high pH*. *Analytical Chemistry*, 2010. **82**(4): p. 1470-1477.
39. Thomsson, K.A., N.G. Karlsson, and G.C. Hansson, *Liquid chromatography-electrospray mass spectrometry as a tool for the analysis of sulfated oligosaccharides from mucin glycoproteins*. *Journal of Chromatography A*, 1999. **854**(1-2): p. 131-139.
40. Prakobphol, A., et al., *Human Low-Molecular-Weight Salivary Mucin Expresses the Sialyl Lewis x Determinant and Has L-Selectin Ligand Activity*. *Biochemistry*, 1998. **37**(14): p. 4916-4927.
41. Xia, B., et al., *Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis*. *Glycobiology*, 2005. **15**(8): p. 747-775.
42. Kawahira, K., et al., *Solid-phase synthesis of O-sulfated glycopeptide by the benzyl-protected glycan strategy*. *Tetrahedron*, 2009. **65**(39): p. 8143-8153.
43. Pabst, M. and F. Altmann, *Influence of Electrosorption, Solvent, Temperature, and Ion Polarity on the Performance of LC-ESI-MS Using Graphitic Carbon for Acidic Oligosaccharides*. *Analytical Chemistry*, 2008. **80**(19): p. 7534-7542.
44. Kawasaki, N., et al., *Structural analysis of sulfated N-linked oligosaccharides in erythropoietin*. *Glycobiology*, 2001. **11**(12): p. 1043-1049.
45. Yagi, H., et al., *Development of structural analysis of sulfated N-glycans by multidimensional high performance liquid chromatography mapping methods*. *Glycobiology*, 2005. **15**(10): p. 1051-1060.
46. von Witzendorff, D., et al., *Characterization of the acidic N-linked glycans of the zona pellucida of prepuberal pigs by a mass spectrometric approach*. *Carbohydrate Research*, 2009. **344**(12): p. 1541-1549.
47. Murakami, T., et al., *Structure determination of a sulfated N-glycans, candidate for a precursor of the selectin ligand in bovine lung*. *Glycoconjugate Journal*, 2007. **24**(4-5): p. 195-206.

48. Lei, M., Y. Mechref, and M.V. Novotny, *Structural Analysis of Sulfated Glycans by Sequential Double-Permethylation Using Methyl Iodide and Deuteromethyl Iodide*. Journal of the American Society for Mass Spectrometry, 2009. **20**(9): p. 1660-1671.
49. Taguchi, T., et al., *Occurrence and structural analysis of highly sulfated multiantennary N-linked glycan chains derived from a fertilization-associated carbohydrate-rich glycoprotein in unfertilized eggs of Tribolodon hakonensis*. European Journal of Biochemistry, 1996. **238**(2): p. 357-367.
50. Damen, C.W.N., et al., *Electrospray Ionization Quadrupole Ion-Mobility Time-of-Flight Mass Spectrometry as a Tool to Distinguish the Lot-to-Lot Heterogeneity in N-Glycosylation Profile of the Therapeutic Monoclonal Antibody Trastuzumab*. Journal of the American Society for Mass Spectrometry, 2009. **20**(11): p. 2021-2033.
51. Zhang, Y., et al., *Distinguishing Phosphorylation and Sulfation in Carbohydrates and Glycoproteins Using Ion-Pairing and Mass Spectrometry*. Journal of the American Society for Mass Spectrometry, 2006. **17**(9): p. 1282-1288.
52. Irungu, J., et al., *Method for characterizing sulfated glycoproteins in a glycosylation site-specific fashion, using ion pairing and tandem mass spectrometry*. Analytical Chemistry, 2006. **78**(4): p. 1181-1190.
53. Jiang, H., J. Irungu, and H. Desaire, *Enhanced detection of sulfated glycosylation sites in glycoproteins*. Journal of the American Society for Mass Spectrometry, 2005. **16**(3): p. 340-348.
54. Toyoda, M., H. Narimatsu, and A. Kameyama, *Enrichment Method of Sulfated Glycopeptides by a Sulfate Emerging and Ion Exchange Chromatography*. Analytical Chemistry, 2009. **81**(15): p. 6140-6147.
55. Wakabayashi, H., et al., *Novel proteoglycan linkage tetrasaccharides of human urinary soluble thrombomodulin, SO4-3GlcA beta 1-3Gal beta 1-3(+/- Sia alpha 2-6)Gal beta 1-4Xyl*. Journal of Biological Chemistry, 1999. **274**(9): p. 5436-5442.
56. Minamisawa, T. and J. Hirabayashi, *Fragmentations of isomeric sulfated monosaccharides using electrospray ion trap mass spectrometry*. Rapid Communications in Mass Spectrometry, 2005. **19**(13): p. 1788-1796.
57. Leach Iii, F.E., et al., *Evaluation of the experimental parameters which control electron detachment dissociation, and their effect on the fragmentation efficiency of glycosaminoglycan carbohydrates*. International Journal of Mass Spectrometry, 2008. **276**(2-3): p. 110-115.
58. Estrella, R.P., et al., *Graphitized carbon LC-MS characterization of the chondroitin sulfate oligosaccharides of aggrecan*. Anal Chem, 2007. **79**(10): p. 3597-606.
59. Staples, G.O., et al., *A chip-based amide-HILIC LC/MS platform for glycosaminoglycan glycomics profiling*. Proteomics, 2009. **9**(3): p. 686-695.
60. Saad, O.M., et al., *Compositional profiling of heparin/heparan sulfate using mass spectrometry: assay for specificity of a novel extracellular human endosulfatase*. Glycobiology, 2005. **15**(8): p. 818-826.
61. Wolff, J.J., et al., *Electron capture dissociation, electron detachment dissociation and infrared multiphoton dissociation of sucrose octasulfate*. European Journal of Mass Spectrometry, 2009. **15**(2): p. 275-281.
62. Wolff, J.J., et al., *Electron Detachment Dissociation of Dermatan Sulfate Oligosaccharides*. Journal of the American Society for Mass Spectrometry, 2008. **19**(2): p. 294-304.
63. Zaia, J., et al., *The Role of Mobile Protons in Negative Ion CID of Oligosaccharides*. Journal of the American Society for Mass Spectrometry, 2007. **18**(5): p. 952-960.
64. Imami, K., Y. Ishihama, and S. Terabe, *On-line selective enrichment and ion-pair reaction for structural determination of sulfated glycopeptides by capillary electrophoresis-mass spectrometry*. Journal of Chromatography A, 2008. **1194**(2): p. 237-242.
65. Zhang, Y., et al., *A novel mass spectrometric method to distinguish isobaric monosaccharides that are phosphorylated or sulfated using ion-pairing reagents*. Journal of the American Society for Mass Spectrometry, 2005. **16**(11): p. 1827-1839.
66. Ohara, K., et al., *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometric Analysis of Polysulfated-Derived Oligosaccharides Using Pyrenemethylguanidine*. Journal of the American Society for Mass Spectrometry, 2009. **20**(1): p. 131-137.
67. Fukuyama, Y., et al., *Ionic Liquid Matrixes Optimized for MALDI-MS of Sulfated/Sialylated/Neutral Oligosaccharides and Glycopeptides*. Analytical Chemistry, 2008. **80**(6): p. 2171-2179.
68. Tissot, B., et al., *Towards GAG glycomics: Analysis of highly sulfated heparins by MALDI-TOF mass spectrometry*. Glycobiology, 2007. **17**(9): p. 972-982.
69. Przybylski, C., et al., *HABA-based ionic liquid matrices for UV-MALDI-MS analysis of heparin and heparan sulfate oligosaccharides*. Glycobiology, 2010. **20**(2): p. 224-234.
70. Garenaux, E., et al., *A single step method for purification of sulfated oligosaccharides*. Glycoconjugate Journal, 2008. **25**(9): p. 903-915.

71. Sangadala, S., U.R. Bhat, and J. Mendicino, *Structures of sulfated oligosaccharides in human trachea mucin glycoproteins*. Mol. Cell. Biochem., 1993. **126**(1): p. 37-47.
72. Lei, M., M.V. Novotny, and Y. Mechref, *Sequential enrichment of sulfated glycans by strong anion-exchange chromatography prior to mass spectrometric measurements*. J. Am. Soc. Mass Spectrom., 2010. **21**(3): p. 348-357.
73. Larsen, M.R., et al., *Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns*. Molecular and Cellular Proteomics, 2005. **4**(7): p. 873-886.
74. Karlsson, N.G., H. Karlsson, and G.C. Hansson, *Strategy for the investigation of O-linked oligosaccharides from mucins based on the separation into neutral, sialic acid- and sulfate-containing species*. Glycoconjugate Journal, 1995. **12**(1): p. 69-76.
75. 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.
76. Mank, M., B. Stahl, and G. Boehm, *2,5-Dihydroxybenzoic Acid Butylamine and Other Ionic Liquid Matrixes for Enhanced MALDI-MS Analysis of Biomolecules*. Analytical Chemistry, 2004. **76**(10): p. 2938-2950.
77. Chen, P., A.G. Baker, and M.V. Novotny, *The use of osazones as matrices for the matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates*. Analytical Biochemistry, 1997. **244**(1): p. 144-151.
78. Mank, M., B. Stahl, and G.n. Boehm, *2,5-Dihydroxybenzoic Acid Butylamine and Other Ionic Liquid Matrixes for Enhanced MALDI-MS Analysis of Biomolecules*. Analytical Chemistry, 2004. **76**(10): p. 2938-2950.
79. Geyer, H., et al., *Core structures of polysialylated glycans present in neural cell adhesion molecule from newborn mouse brain*. Eur J Biochem, 2001. **268**(24): p. 6587-99.
80. Yu, S.-Y., et al., *Enabling techniques and strategic workflow for sulfoglycomics based on mass spectrometry mapping and sequencing of permethylated sulfated glycans*. Glycobiology, 2009. **19**(10): p. 1136-1149.
81. Laremore, T.N., et al., *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometric Analysis of Uncomplexed Highly Sulfated Oligosaccharides Using Ionic Liquid Matrices*. Analytical Chemistry, 2006. **78**(6): p. 1774-1779.
82. Bourne, E.J., et al., *Studies on trifluoroacetic acid. Part I. Trifluoroacetic anhydride as a promotor of ester formation between hydroxy-compounds and carboxylic acids*. J. Chem. Soc., 1949: p. 2976-2979.
83. Stellner, K., H. Saito, and S.I. Hakomori, *DETERMINATION OF AMINOSUGAR LINKAGES IN GLYCOLIPIDS BY METHYLATION - AMINOSUGAR LINKAGES OF CERAMIDE PENTASACCHARIDES OF RABBIT ERYTHROCYTES AND OF FORSSMAN ANTIGEN*. Archives of Biochemistry and Biophysics, 1973. **155**(2): p. 464-472.
84. Hakomori, S., *A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide*. Journal of Biochemistry, 1964. **55**: p. 205-208.
85. Ciucanu, I. and F. Kerek, *A simple and rapid method for the permethylation of carbohydrates*. Carbohydrate Research, 1984. **131**: p. 209-217.
86. Wuhler, M., A.R.d. Boer, and A.M. Deelder, *Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry*. Mass Spectrometry Reviews, 2009. **28**(2): p. 192-206.
87. Ruhaak, L.R., A.M. Deelder, and M. Wuhler, *Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry*. Analytical and Bioanalytical Chemistry, 2009. **394**(1): p. 163-74.
88. Thomsson, K.A., H. Karlsson, and G.C. Hansson, *Sequencing of sulfated oligosaccharides from mucins by liquid chromatography and electrospray ionization tandem mass spectrometry*. Anal Chem, 2000. **72**(19): p. 4543-9.
89. Schulz, B.L., et al., *Identification of two highly sialylated human tear-fluid DMBT1 isoforms: the major high-molecular-mass glycoproteins in human tears*. Biochemical Journal, 2002. **366**(Pt 2): p. 511-520.
90. Schulz, B.L., N.H. Packer, and N.G. Karlsson, *Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis*. Analytical Chemistry, 2002. **74**(23): p. 6088-6097.
91. Staples, G.O., et al., *Improved hydrophilic interaction chromatography LC/MS of heparinoids using a chip with postcolumn makeup flow*. Analytical Chemistry, 2010. **82**(2): p. 516-522.
92. Williams, T.I., et al., *Investigations with O-linked protein glycosylations by matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry*. Journal of Mass Spectrometry, 2008. **43**(9): p. 1215-1223.
93. Adams, J., *Charge-remote fragmentations: analytical applications and fundamental studies*. Mass Spectrom. Rev., 1990. **9**: p. 141-186.

94. von Witzendorff, D., et al., *Characterization of the acidic N-linked glycans of the zona pellucida of prepuberal pigs by a mass spectrometric approach*. Carbohydr Res, 2009. **344**(12): p. 1541-9.
95. Kuster, B., et al., *Sequencing of N-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography*. Analytical Biochemistry, 1997. **250**(1): p. 82-101.
96. Harvey, D.J., et al., *"Internal residue loss": rearrangements occurring during the fragmentation of carbohydrates derivatized at the reducing terminus*. Analytical Chemistry, 2002. **74**(4): p. 734-740.
97. Zhang, J., et al., *Infrared multiphoton dissociation of O-linked mucin-type oligosaccharides*. Anal. Chem., 2005. **77**(1): p. 208-214.

Rapid Commun. Mass Spectrom. 2011, 25, 2611–2618
(wileyonlinelibrary.com) DOI: 10.1002/rcm.5157

Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation

Diarmuid T. Kenny^{1,2}, Samah M. A. Issa^{1,2} and Niclas G. Karlsson^{2*}

¹School of Chemistry, National University of Ireland, Galway, Ireland

²Medical Biochemistry, University of Gothenburg, 405 30 Gothenburg, Sweden

Migration of sulfate groups between hydroxyl groups was identified after collision-induced dissociation (CID) of sulfated oligosaccharides in an ion trap mass spectrometer in negative ion mode. Analysis of various sulfated oligosaccharides showed that this was a common phenomenon and was particularly prominent in sulfated oligosaccharides also containing sialic acid. It was also shown that the level of migration was increased when the sulfate was positioned on the flexible areas of the oligosaccharides not involved in the pyranose ring, such as the extra-cyclic C-6 carbon of hexoses or *N*-acetylhexosamines, or on reduced oligosaccharide. This suggested that migration is dependent on the spatial availability of the sulfate in the ion trap during collision. It is proposed that the migration is initiated when the negatively charged $-\text{SO}_3^-$ residue attached to the oligosaccharide precursor becomes protonated by a CID-induced proton transfer. This is supported by the CID fragmentation of precursor ions depleted of acidic protons such as doubly charged $[\text{M} - 2\text{H}]^{2-}$ ions or the sodiated $[\text{M} + \text{Na} - 2\text{H}]^-$ ions of oligosaccharides containing one sulfate and one sialic acid in the same molecule. Compared to the CID fragmentation of their monocharged $[\text{M} - \text{H}]^-$ ions, no migration was observed in CID of proton depleted precursors. Alternative fragmentation parameters to suppress migration of sulfated oligosaccharides also showed that it was not present when sulfated oligosaccharides were fragmented by HCD (High-Energy C-trap Dissociation) in an Orbitrap mass spectrometer. Copyright © 2011 John Wiley & Sons, Ltd.

Migration of residues in biopolymers during collision-induced dissociation (CID) fragmentation is one of the main problems in structural assignment in mass spectrometry (MS). It has recently been highlighted as a problem for post-translational modification (PTM) identification in phosphoproteomics, where up to 75% of phosphates migrated to a nearby phosphate acceptor, causing ambiguity in site assignment.^[1] Outside the cell, the glycosylation is the dominating PTM, where over half of all secreted and extracellular proteins are glycoproteins.^[2] In glycomics, the phenomenon of migration is constantly a nuisance for the assignment of fucosylation in positive ion mode. The predominate occurrences of fucose migration involve the transfer of fucose from one monosaccharide to an adjacent monosaccharide.^[3] However, Harvey *et al.* also reported long-distance migration on carbohydrates derivatized at the reducing terminus^[4] and Ernst *et al.* reported fucose migration towards the non-reducing end positioned sialic acid in sialyl Lewis x structures.^[5]

In negative ion mode CID, the phenomenon of fucose migration has not been reported. With the informative fragmentation generated in negative ion mode, it appears to be an alternative for analyzing glycosylation without the need for derivatization. Especially for the analysis of sulfated oligosaccharides, this is appealing since the protocol for

advanced derivatization, such as permethylation, is not well established and the negative charge of the sulfate group promotes the ionization of the oligosaccharides in negative ion mode. Here, we report that the migration problem in structural assignment by CID in glycomics is not exclusive for positive ion mode, but can also be prominent in negative ion mode, with migration of sulfate residues occurring on native, reduced and reducing end derivatized oligosaccharides. Sulfation, a secondary modification of oligosaccharides, is an important biological modification since it alters the way glycoproteins interact with other biomolecules. Sulfated glycoprotein oligosaccharides are involved in many biochemical pathways, such as the clearance of thyroid stimulating hormone and luteinizing hormone to lymphocyte homing.^[6–9] However, sulfation is perhaps best known in the study of the glycosaminoglycans, i.e. heparin/heparan sulfate,^[10–12] chondroitin/dermatan sulfate,^[13–16] and keratan sulfate.^[17–19] Sulfation is also commonly found in the *O*-linked oligosaccharides of the mucins lining the respiratory and gastrointestinal tracts^[20] and both sulfated *N*- and *O*-linked oligosaccharides are found in the lymphoid, central and peripheral nerve tissue.^[21]

A mass spectrometer is a rapid and sensitive instrument for the study of sulfated oligosaccharides and has emerged as the analytical tool of choice over traditional methods such as radioactive isotope labeling^[8,22] and nuclear magnetic resonance.^[23,24] However, MS analysis of sulfated oligosaccharides is not without its challenges. In addition to the migration of the sulfate group presented here, the sulfate is notorious for its labile nature in positive and negative

* Correspondence to: N. G. Karlsson, University of Gothenburg, Medical Biochemistry, Box 440, 405 30 Gothenburg, Sweden.
E-mail: niclas.karlsson@medkem.gu.se

ion mode MS analysis. Dissociation of the sulfate group can occur when using matrix-assisted laser desorption/ionization (MALDI) with traditional matrices,^[25] whereas electrospray ionization (ESI) appears to be a softer ionization method and can avoid this problem. Sulfated oligosaccharides are typically in low abundance in complex oligosaccharide samples and liquid chromatography (LC) capable of separating oligosaccharides will help to target and sequence sulfated oligosaccharides by CID in the presence of high abundant neutral and/or sialylated oligosaccharides. ESI-MS in conjunction with LC provides a robust and sensitive method for the analysis of native sulfated oligosaccharides.^[12,15,16,26–32]

Interpretation of tandem mass (MS^2) spectra from CID fragmentation can provide full or partial assignment of the oligosaccharide structure as well as sulfate location. The location of the sulfate has been suggested to be indicated by fragment ions in the low-mass region of the MS^2 spectra corresponding to a particular sulfated monosaccharide residue and even sulfate position.^[31] The presence of fragment ions corresponding to more than one sulfated monosaccharide residue would then indicate the presence of isomeric compounds in the sample. In this report it is shown that this could also be due to experimental artifacts introduced by the CID fragmentation in ion trap MS, inducing structure-dependent sulfate migration.

EXPERIMENTAL

Materials

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless stated otherwise.

Preparation of oligosaccharides

O-Linked oligosaccharides were released from commercial porcine gastric mucin (Sigma-Aldrich) human salivary MUC5B,^[33] chicken ceca samples prepared as per Tierney *et al.*^[34] (gift from Prof. Steven Carrington, University College Dublin, Ireland) in 0.5 M $NaBH_4$ in 50 mM NaOH. Samples were desalted with 0.1 mL of AG50WX8 cation-exchange beads (Bio-Rad, Hercules, CA, USA) packed in C18 zip tips (Millipore, Billerica, MA). Borate complexes were removed by repeated addition/evaporation with 1% acetic acid in methanol (100 μ L for each addition). The released oligosaccharides were dissolved in water for introduction into the LC/MS system. Synthesized NeuAc α 2-6(HSO $_3$ -6)Gal β 1-3GlcNAc β 1 and (HSO $_3$ -6)Gal β 1-3GlcNAc β 1- with a +69Da ethyl azide (Az) tag and synthesized NeuAc α 2-6Gal β 1-3(HSO $_3$ -6)GlcNAc β 1-, NeuAc α 2-3Gal β 1-3(HSO $_3$ -6)GlcNAc β 1- and Gal β 1-3(HSO $_3$ -6)GlcNAc β 1- with a +43 Da ethylamine (EA) tag were generous gifts from Prof James Paulson, Scripps Institute, CA, USA, and were dissolved in 1 μ g/mL of water. HSO $_3$ -3Gal β 1-3(Fuc α 1-4)GlcNAc, HSO $_3$ -3Gal β 1-4(Fuc α 1-3)GlcNAc, chondroitin-4-sulfate and chondroitin-6-sulfate disaccharides (Dextra, Reading, UK) were dissolved in water to a final concentration of 1 μ g/mL. These were also reduced with 0.5 M $NaBH_4$ and desalted as described above.

LC/MS² and LC/MSⁿ of oligosaccharides by CID and HCD

Oligosaccharides were analyzed by LC/MS²^[30] using a 20 cm 250 μ m i.d. column containing 5 μ m porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA, USA) prepared in-house.^[35] Oligosaccharides were eluted using a linear gradient from 0–40% acetonitrile over 40 min at a flow rate of 10 μ L/min. The eluted oligosaccharides were detected in an ESI-IT mass spectrometer (LTQ, Thermo Electron Corp., San Jose, CA, USA) in negative ion mode with a spray voltage of 3.5 kV. Air was used as sheath gas and the mass range was set to m/z 380–2000. Ions specified for sulfated oligosaccharide structures were isolated for MS^n fragmentation by CID with the collision energy set to 30%. Sulfated oligosaccharides were also analyzed with an LC-LTQ Orbitrap mass spectrometer (Thermo Electron Corp.). The LC setup was the same as that described for LC/ESI-ITMS. Ions isolated for HCD fragmentation were acquired with a resolution of 30 000 and subjected to collision with the collision energy set to 60% with an activation time of 30ms.

RESULTS

Identification of CID induced migration of sulfate

Determining the correct position of the sulfate is an important piece of information obtainable from mass spectrometric analysis of sulfated oligosaccharides. The low mass range of the MS^2 spectra often provides this information with fragment ions corresponding to sulfated monosaccharides.^[31] It was noticed that fragmenting sulfated *N*-acetylglucosamine standards with sulfation on either the 6 position on Gal (B_1 ion of m/z 241) or the 6 position on GlcNAc (Y_1/B_2 ion of m/z 282) gave, in addition to dominating low molecular mass fragment ions of the predicted sulfated monosaccharide, also low intense fragment ions corresponding to sulfation of the alternate residue. Even after separation, LC/MS² fragments could be detected showing that this was not due to the presence of isomeric impurities (Fig. 1). This triggered the investigation whether the origin of these fragments was due to a migration process of the sulfate induced by CID and initiated the search for examples of this phenomenon in order to explore the mechanism of this migration, looking at various sulfated glycoconjugates including reducing oligosaccharide structures with a reducing end tag and alditols. It was found that in the collection of accumulated CID spectra of sulfated oligosaccharide alditols in negative ion mode (e.g. in the Uni-Carb-DB database^[36]) that the migration was often negligible. This is exemplified in Fig. 2(A) with the fragment spectra of the monosulfated hexasaccharide Fuc α 1-2Gal β 1-3(HSO $_3$ -6GlcNAc β 1-6)GalNAcol. In this spectrum, only the B_1 fragment ions of m/z 282 and the C_1 fragment ion of m/z 300 corresponding to a sulfated GlcNAc residue were present. In other instances, the situation resembled the sulfated *N*-acetylglucosamine standards (Fig. 1), where the sulfate migration was indicated to be low (<10% compared to the intensity of intact sulfate monosaccharides in the MS^2 spectra). This is exemplified by the fragmentation of the $[M - H]^-$ ion of m/z 610 of a 3-sulfated Lewis x alditol (HSO $_3$ -3Gal β 1-3(Fuc α 1-4)GlcNAcol) (Fig. 2(B)). The B_2 fragment ion of m/z 241 and the C_2 fragment ion of m/z 259 are the fragments

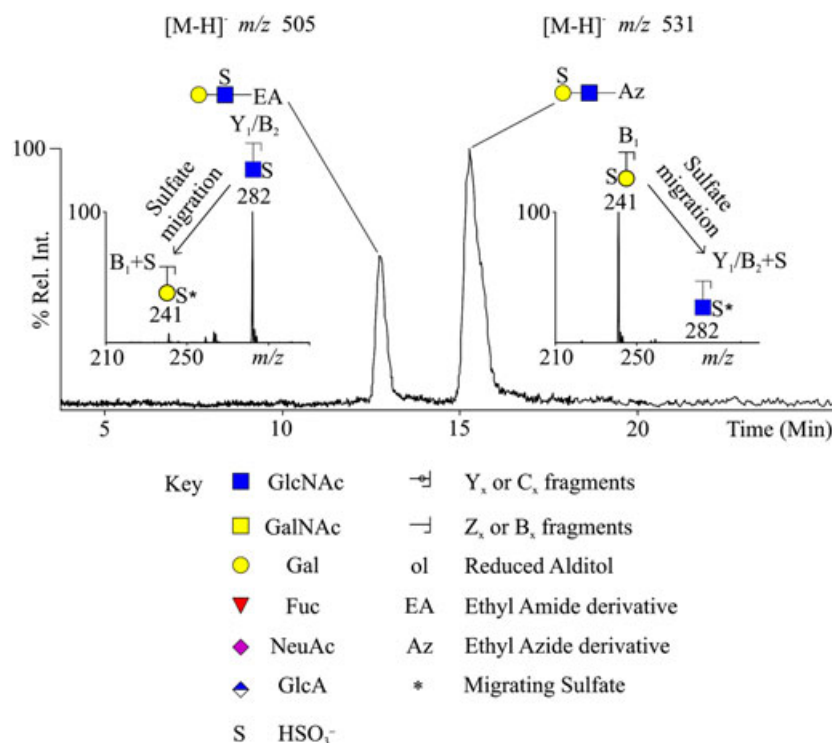


Figure 1. CID-induced migration of sulfate in sulfated *N*-acetylglucosamines in LC/MS². Base peak chromatogram of HSO₃-6Galβ1-4GlcNAcβ1-Az and Galβ1-4(HSO₃-6)GlcNAcβ1-EA separated by LC using a PGC column. Precursor ion was selected for MS² fragmentation by CID. The low-mass regions of the CID spectra are displayed to show fragments from sulfated monosaccharides GlcNAc (*m/z* 282) and Gal (*m/z* 241).

of the non-migrated sulfate on the Gal residue, whereas the presence of low intense fragment ions of *m/z* 284 and 302 correspond to the Z₁ and Y₁ ions with the addition of 80 Da (sulfate). These fragment ions show that the sulfate has migrated to the reducing end GlcNAcol residue. The most striking examples of CID-induced sulfate migration were found when structures containing both sialic acid and sulfate were fragmented. In the CID spectrum (Fig. 2(C)) of the [M - H]⁻

ion of *m/z* 755 from the structure NeuAcα2-6(HSO₃-Galβ1-3)GalNAcol, the presence of an intense fragment ion at *m/z* 370 shows that the migration is a major event, where the sulfate migrates quite readily from the Gal residue to the sialic acid (B_{1α} ion of *m/z* 290) giving the addition of 80 Da (sulfate). This migration is so prominent so it makes it difficult to even detect fragment ions of the original sulfated Gal residue (B_{1β} and C_{1β} of *m/z* 241 and 259, respectively). We subsequently

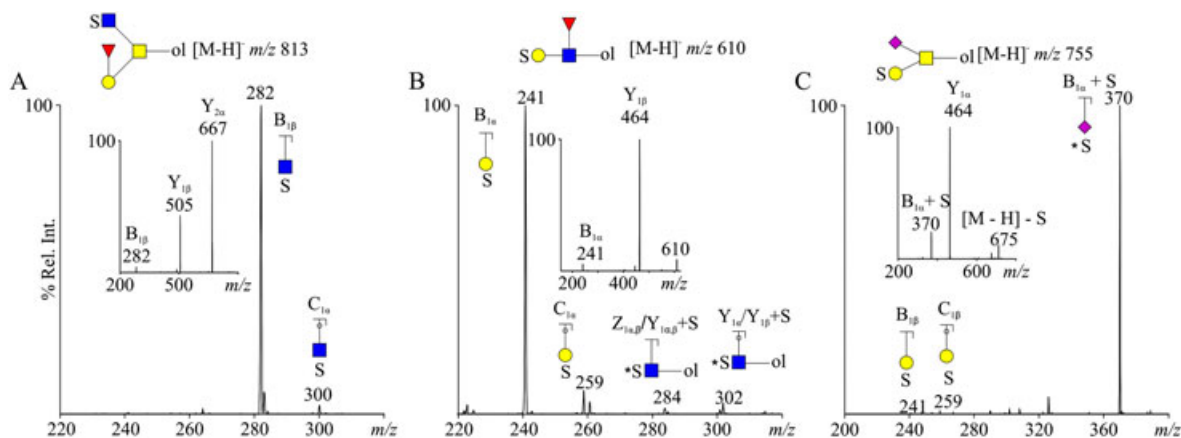


Figure 2. CID-induced sulfate migration in oligosaccharide alditols. The MS² spectrum shows no migration in the sulfated core 1 fucosylated structure Fucα1-2Galβ1-3(HSO₃-6GlcNAcβ1-6)GalNAcol (A), low migration in the sulfated Lewis x structure HSO₃-3Galβ1-3(Fucα1-4)GlcNAcol (B), and high in the sulfated sialylated core 1 oligosaccharide NeuAcα2-6(HSO₃-Galβ1-3)GalNAcol (C). Insert shows the full MS² spectrum from each structure. For key to symbols, see Fig. 1.

screened a whole array of sulfated oligosaccharides, as summarized in Fig. 3. In all our analyzed sulfated oligosaccharides we found that migration of sulfate to sialic acids was always a favorable fragmentation pathway (>10%) in CID, indicating that the structure and/or charge of the sialic acid promote the process of migration. The level of migration on non-sialylated oligosaccharides was usually much lower (0–10%).

Migration of sulfate is due to structure

With the finding that sialic acid plays such a crucial part in promoting the sulfate migration (Fig. 2(C)), it was hypothesized that other carboxylic acid containing residues could have a similar influence on rearrangement events during CID fragmentation. The CID spectra of the $[M - H]^-$ ions of m/z 476 of the two isomeric chondroitin sulfate derived disaccharides [GlcA β 1-3(HSO₃-4)GalNAc and GlcA β 1-3(HSO₃-6)GalNAc] showed that in both these structures the migration of sulfate to the GlcA residue was negligible (6-sulfate isomer) to almost non-existent (4-sulfate isomer) (Fig. 4). This indicated that it was not the carboxylic acid per se that promoted the migration of sulfate in sulfo-sialyloligosaccharides. Further insight into the migration process was provided with the finding that after reduction the CS 4-sulfate disaccharide gave increased CID migration of sulfate to the GlcA residue, whereas the reduction of the 6-sulfate isomer gave virtually

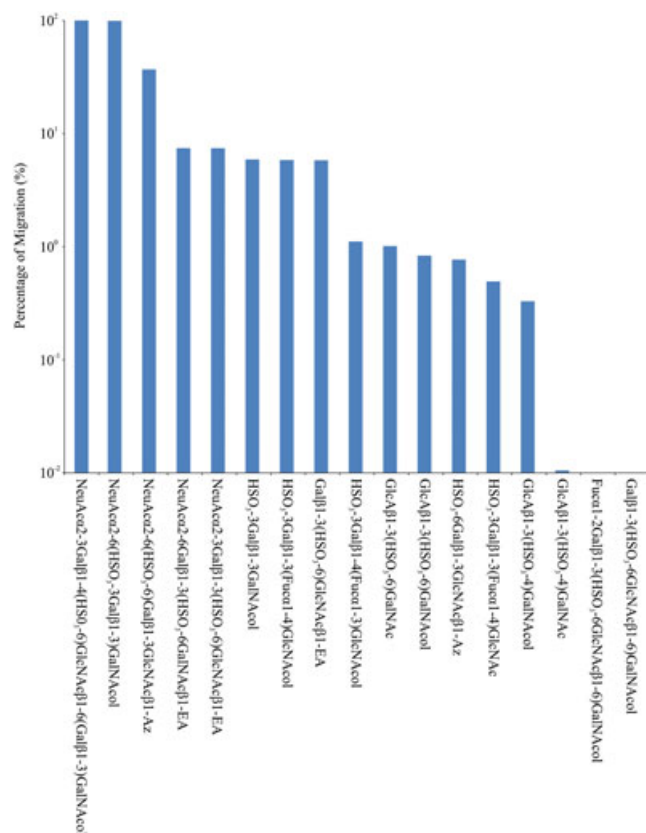


Figure 3. CID-induced sulfate migration in various sulfated oligosaccharides. Bar graph showing the level of migration in different sulfated oligosaccharides. Migration was determined as an ion intensity percentage of the combined ion intensities of fragment ions from sulfated monosaccharide detected in the ion trap.

no altered level of migration (Fig. 4). This indicated that the migration process was a complex rearrangement process depending on the structure of the sulfated oligosaccharide. The data in Fig. 3 indicate that several features of the structures determine the level of migration.

Migration of sulfate is promoted by mobile protons

The presence of mobile protons in negative ion mode CID of oligosaccharides has been shown to significantly influence the quality of CID spectra.^[37,38] $[M - H]^-$ ions of singly sulfated oligosaccharides with one sialic acid or one glucuronic acid residue contain one acidic proton since the number of charges is lower than the number of acidic residues. These acidic protons are probably distributed between the acidic residues containing carboxylic acid and the sulfate. The acidic protons in the precursor ions may be involved in the process of the migration of sulfate by CID. By removing this proton in the precursor ion by selecting the $[M + Na - 2H]^-$ ion of m/z 818 for collision of the NeuAc α 2-6Gal β 1-4(HSO₃-6)GlcNAc β 1-EA it could be shown that the migration of sulfate is reduced, to the level where it could not be detected (Fig. 5). This indicated that mobile protons located to a sulfate residue ($-SO_3H$) are participating in the sulfate rearrangement. The situation with a $[M + Na - 2H]^-$ precursor ion of sulfated sialylated oligosaccharides is similar to the situation of the $[M - H]^-$ ions of sulfated oligosaccharides without sialic acid, in that there is a lack of acidic protons. As a consequence, the amount of sulfate present as ($-SO_3H$) is low and the migration of the non-sialylated sulfated oligosaccharides compared to sialylated ones is reduced (Fig. 3). However, low intensity fragment ions not containing the sulfate group can always be detected also in the CID spectra of structures containing only sulfate as the acidic residue. This indicates that uncharged sulfate in the form of $-SO_3H$ can be formed in the CID process even without an additional acidic group, but instead migrating from hydroxyls on the oligosaccharides. A mechanism for proton migration to sulfate and subsequent fragmentation has been suggested.^[38] The data here indicate an alternative pathway after proton migration, where instead of fragmentation in the CID process the protonated sulfate residue migrates.

Sulfation migration is promoted by hydroxyl groups with high degree of freedom in the precursor ion

Analysis of the results from the chondroitin sulfate disaccharides (Fig. 4) indicated that migration is increased when the sulfate is positioned on the freely available extended arm of the C-6 rather than when positioned on the more restricted C-4 carbon within the rigid ring structure. The level of migration of the C-4 sulfate isomer was shown to increase by increasing the level of conformational freedom of this isomer by reduction, converting the reducing end ring form of the aldose into an open-chain alditol (Figs. 3 and 4). With the C-6 chondroitin sulfate isomer, the migration was already significant even without reduction. These data indicate that it is not the only the degree of rotational freedom of the hydroxyl that receives the migrating sulfate that is important, but also the degree of rotational freedom of the site of sulfate donation. These results support the hypothesis that migration is influenced by the position of the sulfate on a particular

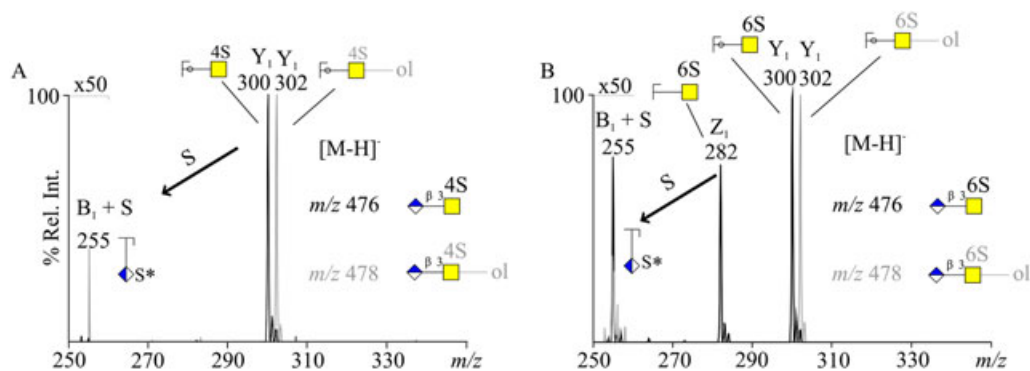


Figure 4. CID fragmentation of chondroitin sulfate disaccharide aldoses and alditols. The CID spectra of the $[M - H]^-$ ion monosulfated chondroitin-4-sulfate (A) and chondroitin-6-sulfate disaccharides (B) showing migration of the sulfate from the GalNAc residue to the GlcA. The spectra of the aldose (grey) and alditol (black) derivatives. For key to symbols, see Fig. 1.

residue of an oligosaccharide and that the more sterically 'flexible' of the sulfate oligosaccharide molecules are increasingly more likely to support migration of sulfate. This is what can be identified in Fig. 3, with migration of sulfate to the flexible side chain of sialic acid and to the flexible linear form of a reduced monosaccharide residue on the reducing end of an oligosaccharide chain.

With the indication that the tertiary structure of the precursor ion of the sulfated oligosaccharide is involved in promoting CID-induced sulfate migration, the process of the rearrangement was further investigated by identification of the location of the sulfate after its migration. The MS^3 spectrum of the $Y_1/B_2 + S$ fragment ion of m/z 282 (sulfate migrating to GlcNAc) from the $[M - H]^-$ ion of $HSO_3-6Gal\beta 1-4GlcNAc\beta 1-EA$ (Fig. 1) showed an intense diagnostic fragment of m/z 139 (Fig. 6(A)). This type of $^{0,4}A$ fragmentation has previously been shown to be a diagnostic fragment for C-6-substituted Hex or HexNAc.^[32,39] An intense $^{0,4}A$ fragment of m/z 139 was also detected in the reverse MS^3 experiment of the $B_1 + S$ fragment ion of m/z 241 (sulfate migrating to Gal) from the $[M - H]^-$ ion of $Gal\beta 1-4(HSO_3-6)GlcNAc\beta 1-Az$ (Fig. 6(B)). These data further strengthen the argument of CID-induced migration of sulfate between less restricted hydroxyls, in these cases between the C-6 positions of adjacent Gal and GlcNAc residues. These data also give further insight into

the particular feature of increased migration of the structure containing both sialic acid and sulfate. The extended flexible glycerol side chain of sialic acid would be an excellent receiving area for a migrating sulfate. The MS^3 spectrum of the $B_1 + S$ fragment ion of m/z 370 from $NeuAc\alpha 2-6(HSO_3-Gal\beta 1-3)GalNAcol$ (Fig. 2(C)) showed in addition to the prominent ions of sialic acid (B_1 ion of m/z 290) also lower intense fragments locating the sulfate to the glycerol side chain, with fragments cleaving between C-6 and C-7 and between C-7 and C-8, corresponding to m/z 139 and 169, respectively (Fig. 6(C)). Whether all hydroxyls of the glycerol chain are the target for migration or only one (e.g. C-9) could not be determined.

Selection of MS fragmentation conditions to prevent sulfate migration

The addition of a sodium salt to generate Na^+ adducts (Fig. 5) as a concept for preventing migration of sulfate in oligosaccharides containing both sulfate and sialic acid is not without problems. In ESI-MS the addition of non-volatile salts is at its best only lowering the ionization efficiency and at its worst being detrimental for the instrument. Selection of multiply charged ions as precursor ions is an attractive alternative to remove these acidic protons in sulfo-sialooligosaccharides

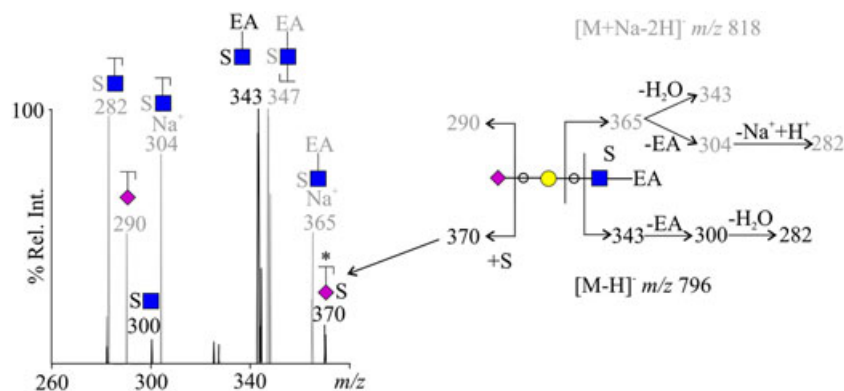


Figure 5. Inhibition of sulfate migration in CID in sulfo-sialooligosaccharides. Comparison of the level of CID-induced sulfate migration of the $NeuAc\alpha 2-6Gal\beta 1-4(HSO_3-6)GlcNAc\beta 1-EA$ oligosaccharide showing the MS^2 spectra of the $[M - H]^-$ ion (black) and the $[M + Na - 2H]^-$ ion (grey) in the low-mass region. For key to symbols, see Fig. 1.

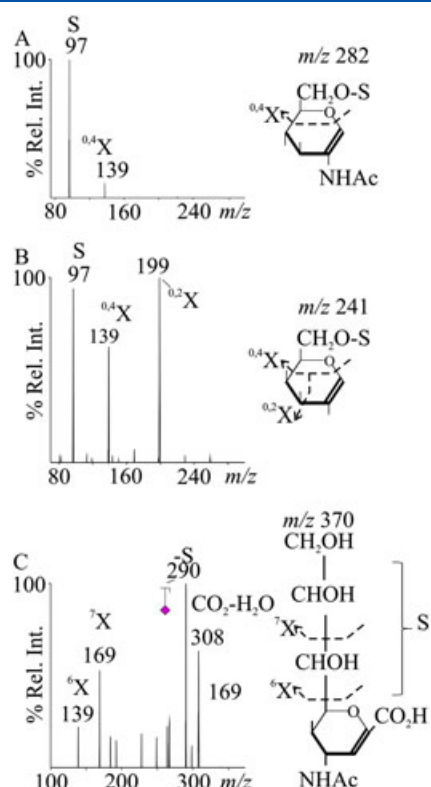


Figure 6. MS³ spectra to identify the target hydroxyl site for CID-induced migration of sulfate. (A) MS³ spectrum of the *m/z* 282 Y₁/B₂ + S fragment ion of sulfate migrating from HSO₃-6Gal to GlcNAc (Fig. 1). (B) MS³ spectrum of the *m/z* 241 B₁ + S fragment ion of sulfate migrating from HSO₃-6GlcNAc to Gal (Fig. 1). MS³ spectrum of the *m/z* 370 B₁ + S fragment ion of sulfate migrating from the 3 position of Gal to NeuAc (Fig. 2(C)). The right-hand side shows proposed sulfation attachment after migration and the supporting fragment pathway.

and prevent migration. Also, with two charges available on a sulfo-sialylated structure (one from the sialic acid carboxyl group and one from the sulfate), the doubly charged species is easy to generate in the mass spectrometer. The singly and doubly charged species of NeuAc α 2-3Gal β 1-4(HSO₃-6GlcNAc β 1-6(Gal β 1-3)GalNAcol ([M - H]⁻ ion of *m/z* 1120 and [M - 2H]²⁻ ion of *m/z* 560, respectively) were isolated for fragmentation by CID (Fig. 7). The presence of a B_{1 α} + S fragment ion of the *m/z* 370 fragment ion in the singly charged precursor indicates migration of the sulfate to the NeuAc residue. Fragmentation of the doubly charged precursor did not show any evidence of sulfate migration (Fig. 7(B)), and the spectrum quality was improved compared to the singly charged CID spectrum, showing more sequence ions.

For structures only containing one sulfate group as the only negative residue, the approach of using multicharged negative precursors in CID is not possible. Since CID in the ion trap is on a time frame where rearrangement is quite frequently observed, we investigated Higher-energy C-trap dissociation (HCD) in an Orbitrap as a possible alternative to CID for fragmentation of sulfated oligosaccharides in an effort to overcome the difficulties associated with sulfate migration. Using an Orbitrap would also provide increased resolution and accuracy to identify larger oligosaccharides and would be an excellent

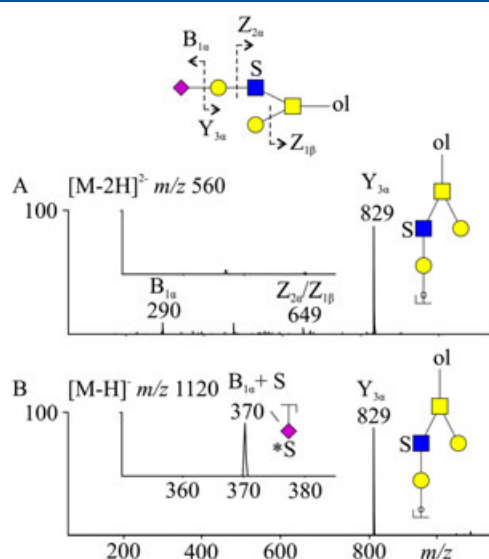


Figure 7. Comparison of the CID-induced sulfate migration in sulfo-sialooligosaccharides in different charge states. The MS² spectrum after CID of the structure NeuAc α 2-3Gal β 1-4(HSO₃-6GlcNAc β 1-6(Gal β 1-3)GalNAcol of the [M - H]⁻ ion (A) and the [M - 2H]²⁻ ion (B). Insert shows the scaled-up region of the mass range where the B₁ + S fragment ion for sialic acid is detected. For key to symbols, see Fig. 1.

complement to our current approach, where we only can sequence <3000Da oligosaccharides on a regular basis due to the mass accuracy of the ion trap. Optimization of the HCD

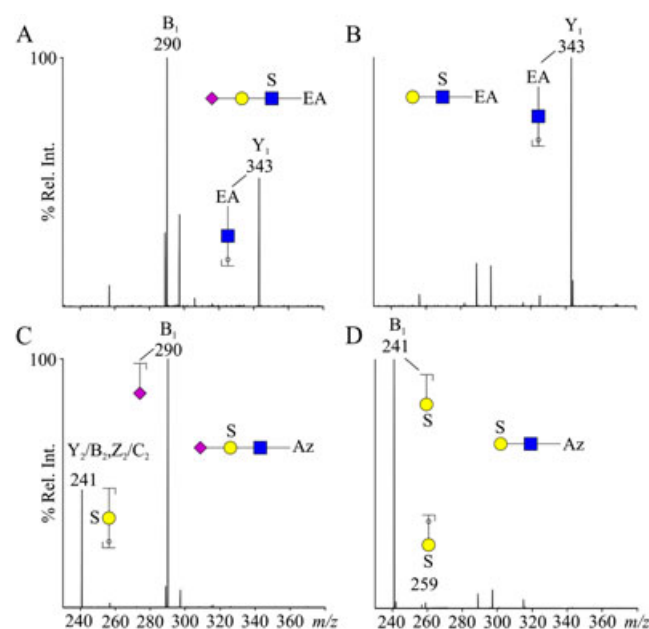


Figure 8. HCD fragmentation of sulfo-sialo-*N*-acetylglucosamine and sulfo-*N*-acetylglucosamine derivatives. Fragment spectra of the [M - H]⁻ ions of (A) NeuAc α 2-3Gal β 1-4(HSO₃-6)GlcNAc β 1-EA and (B) Gal β 1-4(HSO₃-6)GlcNAc β 1-EA with the detection of the sulfate site on the GlcNAc residue (fragment ion of *m/z* 343) and the fragment spectra of the [M - H]⁻ ions of (C) NeuAc α 2-3(HSO₃-6)Gal β 1-4GlcNAc β 1-Az and (D) HSO₃-6Gal β 1-4GlcNAc β 1-Az with the detection of the sulfate site on the Gal residue (fragment ion of *m/z* 241). For key to symbols, see Fig. 1.

showed that a collision energy of 60% provided informative fragmentation of $[M - H]^-$ ions of both unsialylated and sialylated sulfated oligosaccharides. Figure 8 shows the HCD spectra from $[M - H]^-$ ions of both sialylated and non-sialylated *N*-acetylglucosamine derivatives with the sulfate positioned on either the Gal or the GlcNAc. Migration of sulfate, especially towards the sialic acid residue which is the most prominent migration pathway in CID, was completely absent in HCD. In CID, the reducing end tags (EA) are cleaved from the GlcNAc residue (Fig. 1), whereas in the HCD experiment these tags remain so that a modified sulfated GlcNAc shows a Z_1 ion of m/z 343 with an EA tag in both the non-sialylated and sialylated species (Figs. 8(A) and 8(B)). Sulfated Gal was also identified by HCD by the presence of the fragment ion of m/z 241 (B_1/Y_1 in Fig. 8(C) and B_1 in Fig. 8(D), respectively). With the higher radiofrequency voltage in the C-trap that is a feature of HCD, it is possible that the uncharged $-SO_3H$ does not have time to form, thus stabilizing the sulfate from migration. The results suggest that HCD may be a preferential fragmentation pathway for the analysis of sulfated oligosaccharides providing correct, unambiguous assignment of sulfate to a particular monosaccharide residue.

DISCUSSION

The presented data shows the nature of the sulfate migration induced by CID in ITMS. Despite efforts to minimize the migration by increasing or decreasing the activation time in the ion trap, our experience was that it was difficult to control.

Two molecular properties were identified as important for promoting the migration. The first prerequisite was that an acidic proton had to be present to initiate sulfate migration. In structures containing both sialic acid and sulfate, the presence of acidic protons made migration significant in the CID of their $[M - H]^-$ ions but was lower when multiply charged ions were selected for fragmentation. Consequently, the absence of acidic protons in non-sialylated singly sulfated structures made the level of sulfate migration in their CID spectra lower (Fig. 3). A second prerequisite for the sulfate to migrate was that the precursor ion provided the molecular environment to allow migration from one hydroxyl to another. In this context the flexibility of hydroxyls is important to allow them to bridge between the donor and the acceptor site. We propose that the areas of high degrees of conformational freedom within the oligosaccharides are triggering this migration, based on the data presented. This makes the C-6 hydroxyl groups of Hex and HexNAc residues excellent donors and acceptors of a migrating sulfate, as well as the flexible glycerol side chain of sialic acid (Fig. 9). Reduction from aldoses to alditol, thus increasing the flexibility of the reducing end sugar residue by changing it from the rigid ring structure to the increasingly flexible extended structure, also increased sulfate migration in CID. We found that sulfate on the C-6 branch of GalNAcol appeared to be quite stable, while sulfate on the C-3 branch was more prone to migrate to the reducing end GalNAcol. This was probably due to the flexible C-4 to C-6 tail of a GalNAcol to provide spatial closeness to a C-3-branch-linked sulfate (Fig. 9), while C-6-branch-linked sulfate was constrained in this interaction (Fig. 2(A)).

It was identified that the migration was a phenomenon exclusive to the time frame of the CID in an ion trap, where

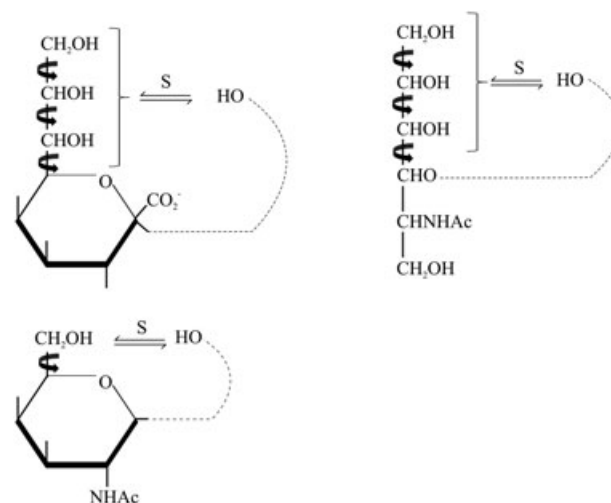


Figure 9. Summary of pathways for migration of sulfate in CID. Proposed migration of sulfate to and from areas of high degree of rotational freedom in oligosaccharides containing sialic acid, alditols and hexoses with a free rotational C-6.

rearrangement has time to occur. The experiment with HCD shows that the sulfate migration can be controlled. We also have indications that CID in a Q-TOF also allows a better control of the sulfate migration.

Analysis of non-derivatized oligosaccharides in negative ion mode MS is an appealing concept, especially for sulfated oligosaccharides where derivatization can be problematic due to the labile nature of the sulfate residue. With the low consumption of valuable samples due to sensitive detection of native oligosaccharides by ESI-MS, there is always the possibility of subsequent analysis by derivatization, such as permethylation or reducing end derivatization, in order to direct fragmentation for further structural analysis. With the ability to control the sulfate migration and to generate informative fragmentation spectra as indicated here, MS^n provides structural information of sulfated oligosaccharides that allow this approach to highlight and address biological questions of the role of sulfated oligosaccharides.

Acknowledgements

Professors James Paulson, Scripps Institute, and Steven Carrington, University College Dublin, are gratefully acknowledged for their gifts of sulfated glycoconjugates. Dr Kristina Thomsson (University of Gothenburg) is acknowledged for her valuable input during the preparation of this manuscript. This work was supported by the Swedish Research Council (621-2010-5322). The mass spectrometers were obtained by grants from the Swedish Research Council (342-2004-4434) and from the Knut and Alice Wallenberg Foundation (KAW2007.0118).

REFERENCES

- [1] J. N. Heck, D. L. Mellman, K. Ling, Y. Sun, M. R. Wagoner, N. J. Schill, R. A. Anderson. *Crit. Rev. Biochem. Mol. Biol.* **2007**, *42*, 15.
- [2] M. Toyoda, H. Narimatsu, A. Kameyama. *Anal. Chem.* **2009**, *81*, 6140.

- [3] M. Wührer, C. A. M. Koeleman, C. H. Hokke, A. M. Deelder. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1747.
- [4] D. J. Harvey, T. S. Mattu, M. R. Wormald, L. Royle, R. A. Dwek, P. M. Rudd. *Anal. Chem.* **2002**, *74*, 734.
- [5] B. Ernst, D. R. Muller, W. J. Richter. *Int. J. Mass Spectrom. Ion Processes* **1997**, *160*, 283.
- [6] H. Kawashima. *Biol. Pharm. Bull.* **2006**, *29*, 2343.
- [7] J. U. Baenziger, S. Kumar, R. M. Brodbeck, P. L. Smith, M. C. Beranek. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 334.
- [8] J. U. Baenziger, E. D. Green. *Biochim. Biophys. Acta* **1988**, *947*, 287.
- [9] Y. Imai, M. S. Singer, C. Fennie, L. A. Lasky, S. D. Rosen. *J. Cell Biol.* **1991**, *113*, 1213.
- [10] N. Jastrebova, M. Vanwildemeersch, U. Lindahl, D. Spillmann. *J. Biol. Chem.* **2010**, *285*, 26842.
- [11] M. Viola, D. Vigetti, E. Karousou, B. Bartolini, A. Genasetti, M. Rizzi, M. Clerici, F. Pallotti, G. De Luca, A. Passi. *Electrophoresis* **2008**, *29*, 3168.
- [12] J. E. Turnbull, R. L. Miller, Y. Ahmed, T. M. Puvirajesinghe, S. E. Guimond. *Methods Enzymol.* **2010**, *480*, 65.
- [13] A. Muresan, M. Galusca, D. G. Seidler, N. Dinca, A. D. Zamfir, in *NATO Advanced Research Workshop on Applications of Mass Spectrometry in Life Safety*, (Eds: C. Popescu, A. D. Zamfir, N. Dinca), Springer, Arad, Romania, **2007**, p. 85.
- [14] J. Zaia, J. E. McClellan, C. E. Costello. *Anal. Chem.* **2001**, *73*, 6030.
- [15] N. Volpi. *Curr. Pharm. Des.* **2006**, *12*, 639.
- [16] J. Zaia, X. Q. Li, S. Y. Chan, C. E. Costello. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 1270.
- [17] N. G. Karlsson, B. L. Schulz, N. H. Packer, J. M. Whitelock. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2005**, *824*, 139.
- [18] Y. Zhang, T. Iwamoto, G. Radke, Y. Kariya, K. Suzuki, A. H. Conrad, J. M. Tomich, G. W. Conrad. *J. Mass Spectrom.* **2008**, *43*, 765.
- [19] T. Oguma, S. Tomatsu, A. M. Montano, O. Okazaki. *Anal. Biochem.* **2007**, *368*, 79.
- [20] S. K. Linden, P. Sutton, N. G. Karlsson, V. Korolik, M. A. McGuckin. *Mucosal Immunol.* **2008**, *1*, 183.
- [21] S.-Y. Yu, S.-W. Wu, H.-H. Hsiao, K.-H. Khoo. *Glycobiology* **2009**, *19*, 1136.
- [22] K. G. Bowman, B. N. Cook, C. L. de Graffenried, C. R. Bertozzi. *Biochemistry* **2001**, *40*, 5382.
- [23] K. Hård, G. Van Zadelhoff, P. Moonen, J. P. Kamerling, F. G. Vliegthart. *Eur. J. Biochem.* **1992**, *209*, 895.
- [24] J. J. M. Van Rooijen, J. P. Kamerling, J. F. G. Vliegthart. *Eur. J. Biochem.* **1998**, *256*, 471.
- [25] Y. Fukuyama, S. Nakaya, Y. Yamazaki, K. Tanaka. *Anal. Chem.* **2008**, *80*, 2171.
- [26] K. A. Thomsson, M. Backstrom, J. M. Holmen Larsson, G. C. Hansson, H. Karlsson. *Anal. Chem.* **2010**, *82*, 1470.
- [27] R. P. Estrella, J. M. Whitelock, N. H. Packer, N. G. Karlsson. *Anal. Chem.* **2007**, *79*, 3597.
- [28] A. Antonopoulos, P. Favetta, W. Helbert, M. Lafosse. *J. Chromatogr. A* **2007**, *1147*, 37.
- [29] M. Pabst, F. Altmann. *Anal. Chem.* **2008**, *80*, 7534.
- [30] H. Karlsson, J. M. Larsson, K. A. Thomsson, I. Hård, M. Bäckström, G. C. Hansson. *Methods Mol. Biol.* **2009**, *534*, 117.
- [31] K. A. Thomsson, N. G. Karlsson, G. C. Hansson. *J. Chromatogr. A* **1999**, *854*, 131.
- [32] K. A. Thomsson, H. Karlsson, G. C. Hansson. *Anal. Chem.* **2000**, *72*, 4543.
- [33] K. A. Thomsson, B. L. Schulz, N. H. Packer, N. G. Karlsson. *Glycobiology* **2005**, *15*, 791.
- [34] J. B. Tierney, E. Matthews, S. D. Carrington, G. Mulcahy. *J. Parasitol.* **2007**, *93*, 634.
- [35] D. T. Kenny, L. Ali, S. Issa, N. G. Karlsson, in *Sample Preparation in Biological Mass Spectrometry*, (1st edn.), (Eds: A. R. Ivanov, A. V. Lazarev), Springer, **2011**, p. 498.
- [36] Available: www.unicarb-DB.com.
- [37] J. L. Seymour, C. E. Costello, J. Zaia. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 844.
- [38] J. Zaia, M. J. Miller, J. L. Seymour, C. E. Costello. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 952.
- [39] N. G. Karlsson, H. Karlsson, G. C. Hansson. *J. Mass Spectrom.* **1996**, *31*, 560.

Semi-Quantitative Data Analysis for Membrane Associated *N*-linked Oligosaccharides Highlights Differences in the Glycosylation of Two Enriched Membranes Samples

Diarmuid T. Kenny^{1,2}, Kristina A. Thomsson¹, Niclas G. Karlsson¹

¹Department of Medical Biochemistry, Göteborg University, Sweden and ²School of Chemistry, National University Ireland Galway, Ireland

Glycomic data is difficult to display due to the fact that the oligosaccharides are diverse, and structural and functional aspects overlap, where an attribute cannot be explained by only one individual structure. We set out to investigate how we could transform glycomic MS data derived from the analysis of *N*-linked oligosaccharides into easily accessible information representing differences in glycosylation. In order to do this we adopted a method for the global enrichment of cellular membrane using carbonate extraction and analysed two samples with differing levels of sialylation. We could identify 41 different oligosaccharides based on their composition. By measuring the MS intensities of the compositions that were identified, we could determine the relative abundance of the individual oligosaccharides. By sorting the identified oligosaccharides by their structural characteristics (high-mannose, high and low sialylation content) and applying mass spectrometric average composition (MSAC) analysis, we could identify particular traits about the enriched membranes. Using this approach it was possible to highlight differences in glycosylation profile in order to make the data from complex MS-spectra accessible for structural comparison.

N-linked | glycosylation | LC-MS | Sialylation

The cellular membrane proteins account for up to 30 percent of the total cellular protein content and are usually found to be *N*-linked glycosylated¹. *N*-linked oligosaccharides are, for instance, present on the plasma membrane where it has been identified as a mediator of extracellular signalling and involved in immunological events². Therefore, characterising the structural features of the oligosaccharides that mediates this signalling is important. Glycomics of the immunologically important cell surface oligosaccharide epitopes can aid in our understanding of their role in immunological signalling. Mass spectrometry (MS) has revolutionised the field of glycomics and has become the primary instrument of choice for a glycomic platform³⁻⁵. When coupled to an LC-system, it is ideally suited to characterising the mixture of oligosaccharides structures typically present on a glycoprotein⁶. The data generated by glycomic investigation are usually quite complex and confusing, since similar structural features are present on several structures and the information about their biological importance is usually quite difficult to digest. The communication between the glyco biologist and the glycoanalyst would be greatly enhanced if they conformed to using a common language and terminology of important epitopes such as high-mannose, sialylation, blood group antigens, sulfation, core types, *N*-acetylglucosamine

elongation. This would aid in the amalgamation of reports and also provide a platform for better communication regarding MS specific information that is not obtainable by other methods such as antibody binding and lectin epitopes including details about antenna distribution that serves as scaffold for the terminal epitopes and potentially semi-quantitative data relating to the abundance of oligosaccharides. In this report we investigate how to display the glycomic information obtained from MS data when comparing the glycosylation from two different mammalian membranes.

Materials and Methods

All materials were obtained from Sigma Aldrich (St Louis, MO) unless otherwise stated. The 18 MΩ water was produced using the MilliQ water purification system (Millipore, Billerica, MA)

Enrichment of Membrane proteins from mammalian cells. Membrane proteins were enriched by precipitation in sodium carbonate^{7,8}. Cellular material (100 mg) of was suspended in 100 mM sodium carbonate. The solution was sonicated on ice at 40% maximum amplitude four times for 15 seconds and cooled on ice in between each sonication for 1 minute. After sonication, the solution was transferred to a beaker and 100 mM of sodium carbonate was added to a final volume of 100 mL and stirred for 5 minutes at 4 °C. Cell debris was removed by centrifugation at 5,000 g for 10 minutes. After ultracentrifugation, the supernatant was transferred to a new centrifuge tubes and ultracentrifuged at 115,000 g for 75 minutes. The supernatant was discarded and the pellet was washed in 50 mM tris(hydroxymethyl)aminomethane-HCl (Tris-HCl) pH 7.3 and ultracentrifuged at 115,000 g for 75 minutes. The wash step was repeated once. The washed membrane containing pellets were solubilised in 7.0 M urea, 2.0 M thiourea, 40 mM Tris base and 4% w/v 3-[(3 cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS).

Release of *N*-linked oligosaccharides for LC-MS analysis. The solubilised membrane extracts were blotted onto Immobilon-P PVDF membranes (Millipore, Billerica, MA) and stained with DB71 and de-stained in 10% acetic acid in 40% ethanol. The membranes were blocked with 1% polyvinylpyrrolidone (PVP) in methanol. The *N*-linked oligosaccharides were enzymatically released from the protein by PNGase F. The enzyme (5 μL, 5 units) was added to each blot and incu-

Correspondence: Niclas G. Karlsson, Dept. Medical Biochemistry, University of Gothenburg, Box 440, 405 30 Gothenburg, SWEDEN; Phone: +46 31 786 6528. Email: Niclas.karlsson@medkem.gu.se

bated for 10 minutes at 37°C. After 10 minutes incubation a further 5 μ L of PNGase F and 10 μ L of water was added and the blots were incubated overnight at 37°C. The released oligosaccharides were reduced by the addition of 0.5 M NaBH₄ in 50 mM NaOH and incubated at 60°C overnight. The reaction was suspended with acetic acid and the samples were desalted with 60 μ L of AG 50W-X8 cation exchange beads (Bio-Rad, Hercules, CA) packed in C18 zip tips (Millipore, Billerica, MA). Borate complexes were removed by repeated addition/evaporation with 1% acetic acid in methanol (100 μ L for each addition). The released oligosaccharides were dissolved in water prior analysis by LC-MS.

LC-MS of released N-linked oligosaccharides. Released membrane associated oligosaccharides were analysed by LC-MS using a 10 cm \times 250 μ m I.D. column containing 5- μ m porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA) prepared in-house⁹. Oligosaccharides were eluted using a linear 8mM aqueous ammonium bicarbonate/acetonitrile gradient from 0-40% in 40 minutes at a flow rate of 10 μ L/minute. The eluted oligosaccharides were analysed using an ESI-IT MS (LTQ, Thermo Electron Corp., San Jose, CA) operating in negative ion mode with a spray voltage of -3.5 kV. Oligosaccharides were detected as [M-H]⁻, [M-2H]²⁻ and [M-3H]³⁻ over a scan range of *m/z* 380-2000. Individual oligosaccharides were isolated for fragmentation by collision induced dissociation (CID) with the collision energy of 35% and a dynamic integration time. The oligosaccharide composition was determined by manual interpretation of the MS and MS/MS spectral data. For relative abundance determination, compositions were identified by their *m/z* value and retention time. An extracted ion chromatogram was generated for each oligosaccharide and the MS intensity for each MS scan was summed to determine the relative abundance of the particular oligosaccharide. The MSAC analysis was performed as described by Hayes et al¹⁰. Briefly, the intensities for each composition identified were determined. The compositions were reduced to their monosaccharide content by determining the number GlcNAc, Man, Gal, Fuc and NeuAc residues. The monosaccharide content for each composition was multiplied by the relative abundance of the particular structure and then each parameter summed over the entire sample and the percentage of each summed monosaccharide was calculated to give the MSAC data. An example of this calculation is given in Supplementary Table 1.

Result

Compositional analysis of released oligosaccharides.

LC-MS has been used extensively for the analysis of oligosaccharides. Here it was utilized for the analysis of N-linked oligosaccharides released from enriched membrane proteins. Two membranes fractions were enriched from whole lysates of a high and low sialic acid (HSA and LSA) containing cells by precipitation with sodium carbonate. The LC-MS (Figure 1), shows that the glycosylation profile between the samples is very different. The first level of interpretation utilized the compositional analysis to reveal that majority of the most abundant structures in HSA contained one or two sialic acid residues, whereas the most abundant structures on the LSA were all unsialylated (Table 1) This analysis verified that the oligosaccharides differed in their amount of sialic acid. Having identified that the samples showed different profiles, the glycosylation pattern was further analysed focussing on

both high and low abundant oligosaccharides. Of the 41 identified oligosaccharide compositions, 21 were present on both membranes and 20 were detected on the LSA membranes only. Compositional analysis of these structures revealed that the oligosaccharides common to both membranes contained hexoses (Hex), N-acetylhexosamines (HexNAcs) with and without deoxyhexoses (interpreted to be fucose (Fuc)) and sialic acids (NeuAc). Additional structures present only in the LSA contained only Hex, HexNAcs with or without Fuc (see Table 1) but no sialic acid.

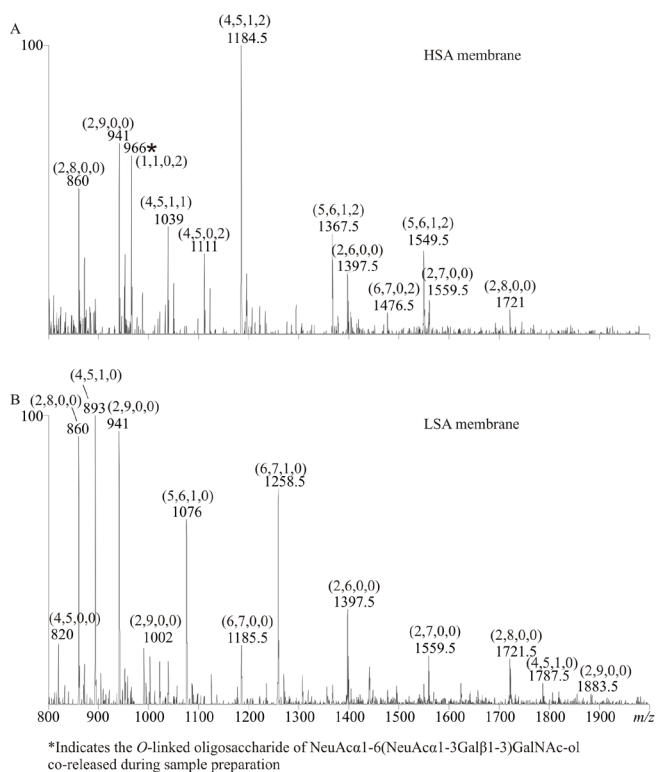


Figure 1: LC-MS of N-linked oligosaccharides released from (a) HSA and (b) LSA membranes. The oligosaccharides were detected as the [M - H]⁻, [M - 2H]²⁻, [M - 3H]³⁻ ions between 20 - 40 min. The composition of the most abundant oligosaccharides is indicated as (Hex, HexNAc, Deoxy-Hex, Sialic Acid).

Identification of high-mannose and complex structures on membranes.

The composition analysis generates a limited amount of information on the glycosylation profile of the membranes as it lacks structural details which are important when investigating the biological role of the membrane glycosylation. However, it is known that all N-linked oligosaccharides share a common GlcNAc₂Man₃ core sequence. This sequence can be elongated by the addition of more Man residues to create high-mannose structures, by the addition of N-acetylglucosamine (GlcNAc, Galβ1-4GlcNAc) residues with the possible further addition of Fuc and/or sialic acid residues to create complex structures or by a combination of both to produce hybrid structures. We can therefore utilise our knowledge of the biosynthetic pathway when assigning the structure of the oligosaccharides to garner structure specific information missing from the compositional analysis. The LC-MS data shows the presence of both high-mannose and complex structures on the membrane proteins (Table 1). The high-mannose structures were detected with 5 (Man5, *m/z* 1235), 6 (Man6, *m/z* 1397.5) and 7 (Man7, *m/z* 1559.5)

Table 1: Composition analysis of the oligosaccharides identified on the High and Low SA membrane. The composition is presented as (HexNAc, Hex, Deoxyhex, Sialic Acid). The structures for each oligosaccharides is presented using CFG annotation.

Oligosaccharides Identified on HSA and LSA				Oligosaccharides Identified on LSA only			
	[M-nH] ⁻ⁿ	Composition	Structure		[M-nH] ⁻ⁿ	Composition	Structure
1	1235 ¹⁻	(2,5,0,0)		22	911 ¹⁻	(2,3,0,0)	
2	1397.5 ¹⁻	(2,6,0,0)		23	1276 ¹⁻	(3,4,0,0)	
3	1559.5 ¹⁻	(2,7,0,0)		24	1480.5 ¹⁻	(4,4,0,0)	
4	1721.5 ¹⁻	(2,8,0,0)		25	1104 ²⁻	(6,6,0,0)	
	860 ²⁻			26	1185.5 ²⁻	(6,7,0,0)	
5	1883.5 ¹⁻	(2,9,0,0)		27	1368 ²⁻	(7,8,0,0)	
	941 ²⁻			28	1550.5 ²⁻	(8,9,0,0)	
6	1039 ²⁻	(4,5,1,1)		29	1422 ¹⁻	(3,4,1,0)	
7	1220 ²⁻	(5,5,1,1)		30	995 ²⁻	(5,5,1,0)	
8	1404 ²⁻	(6,7,1,1)		31	1177 ²⁻	(6,6,1,0)	
9	1111 ²⁻	(4,5,0,2)		32	1360 ²⁻	(8,7,1,0)	
10	1184.5 ²⁻	(4,5,1,2)		33	1441 ²⁻	(7,8,1,0)	
11	1367.5 ²⁻	(5,6,1,2)		34	1542.5 ²⁻	(8,8,1,0)	
12	1476.5 ²⁻	(6,7,0,2)		35	1623.5 ²⁻	(8,9,1,0)	
13	1549.5 ²⁻	(6,7,1,2)		36	1806.5 ²⁻	(9,10,1,0)	
14	1732.5 ²⁻	(7,8,1,2)		37	1988 ²⁻	(10,11,1,0)	
15	1915 ²⁻	(8,9,1,2)		38	1447 ³⁻	(11,12,1,0)	
16	1057 ¹⁻	(2,3,1,0)		39	1568 ³⁻	(12,13,1,0)	
17	1642.5 ¹⁻	(4,5,0,0)		40	1690.5 ³⁻	(13,14,1,0)	
	820 ²⁻			41	1812 ³⁻	(14,15,1,0)	
18	1788.5 ¹⁻	(4,5,1,0)					
	893 ²⁻						
19	1002.5 ²⁻	(5,6,0,0)					
20	1076 ²⁻	(5,6,1,0)					
21	1258.5 ²⁻	(6,7,1,0)					

Man residues as their [M-H]⁻ ions and with 8 (Man8, *m/z* 1721.5 and 860) and 9 (Man9, *m/z* 1883.5 and 941) Man residues detected as the [M-H]⁻ and/or [M-2H]²⁻ ions. The complex structures were detected as the [M-H]⁻, [M-2H]²⁻ and [M-3H]³⁻ ions (Table 1). Of the complex structures identified, 16 were detected on both membranes. These were predominantly core fucosylated with or without terminal sialylation. The additional complex structures identified on the LSA membrane were all unsialylated extended with up to 12 lacNAc residues. The data obtained by interpretation of masses into composition and biosynthetic pathways are providing the first level of interpretation and is required for stringent documentation of the raw data. This is also the minimal amount of data that is required for submission to structural databases or knowledgebases and for recording and future references.

It is then the glycoanalysts obligation to do further interpretation of the data in order to communicate the result to a wider community.

Semi-quantitative display of individual structures with high-mannose and complex type glycosylation. To better reflect the biological aspects of the membrane glycosylation, the oligosaccharides can also be displayed according to their biological characteristics. The lack of quantitative data in Table 1 is limiting the amount of biologically relevant information that can be ascertained from it. Using the predicted oligosaccharide structures from LC-MS and the intensity of each composition, it is possible generate semi-quantitative data that shows there is a significant variation in the abundance of the oligosaccharides between HSA and LSA membranes (Figure 2). Further information about the glycosylation can be obtained by sorting the oligosaccharides into high-mannose and complex type and this reveals that the main difference

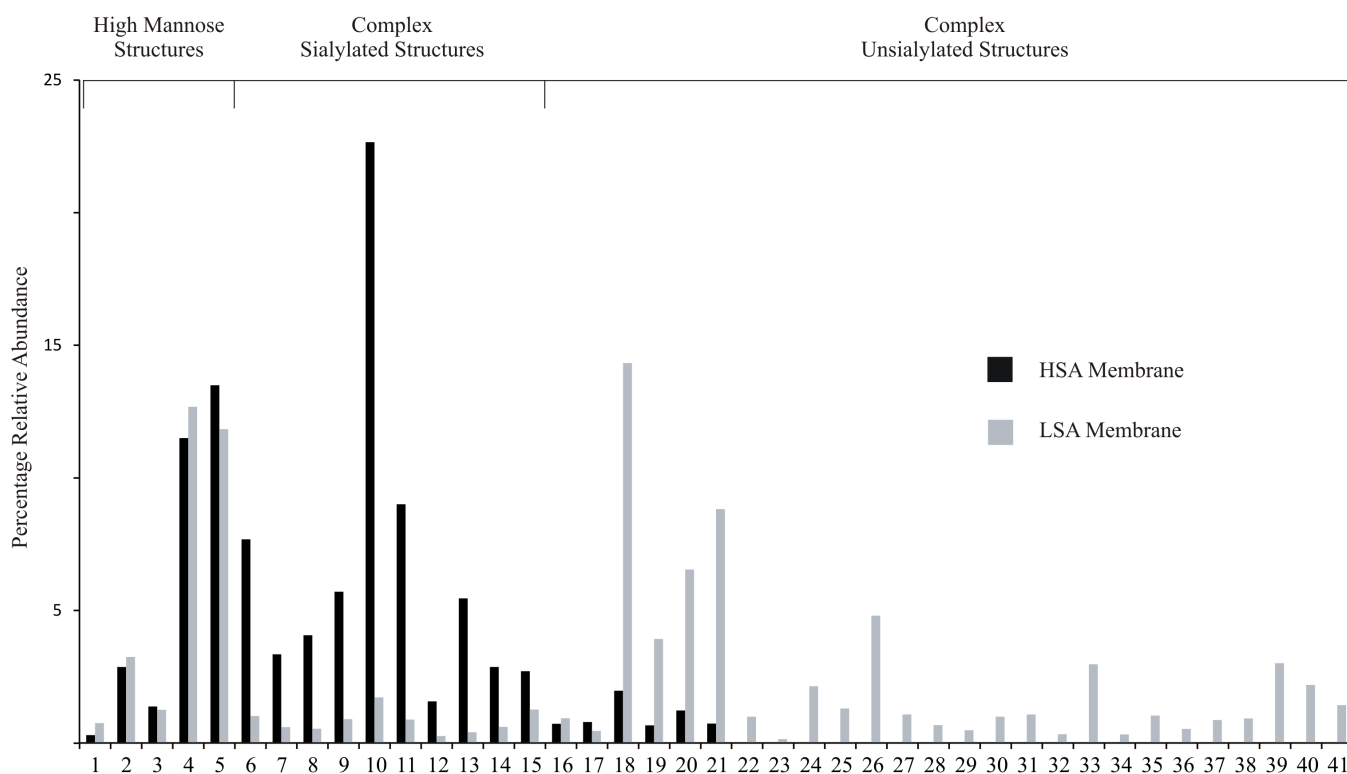


Figure 2: The relative abundance of the *N*-linked oligosaccharides detected on the HSA (back) and LSA (grey) membranes. The high mannose, complex sialylated structures and complex unsialylated structures are indicated. The number of each structure corresponds to their numbering in Table 1.

between the samples lie in the amount of sialylated and unsialylated structures. However, the MS based information is still obscuring the picture by sorting the x-axis into features that are only accessible for a trained glyco biologist, and would be difficult even for researchers in the field to digest and make the connection between the identified structures and their function.

Average based MS data to display glycosylation differences. We have described a method for how to use the structural information from MS and the intensities to calculate a mass spectrometry average composition (MSAC)¹⁰ (Figure 3A). The monosaccharide content of the membranes, based on the semi-quantitative data, can be used to display the glycosylation within a global context. This approach used for displaying *O*-linked semi-quantitative data was found to be superior in highlighting the difference in the amount of *N*-linked sialic acid present on both membranes. Furthermore, from the data we can see a higher amount of HexNAc and Hex in the complex oligosaccharides residues, indicating a higher abundance of lacNAc residues. For *O*-linked data there are a number of cores that could be terminated or extended, but *N*-linked oligosaccharides all have a common core. In figure 3Ai the monosaccharide content was determined with the inclusion of the $\text{GlcNAc}_2\text{Man}_3$ core sequence. However, since all structures contain this core structures, the MS based monosaccharide content data was normalized by excluding the core sequence to stress the differences in the antennas (Fig 3Aii). Combining this MSAC method with info about *N*-linked glycosylation (complex, high-mannose) provides an efficient way to grasp the MS data. In figure 3B we

can see that the abundance of high-mannose and complex structures for both membranes is similar and that the majority of the oligosaccharides in both samples are complex structures. Of those complex structures, the sialylated structures are the major component on the HSA membrane, whereas they are only a minor component on the LSA membrane (Figure 3C). The relative abundance of the individual high-mannose oligosaccharides did not vary greatly between the samples, the complex oligosaccharides differed significantly with the sialylated structures having greater abundance on HSA membranes than LSA membranes. Conversely the most abundance structures on the LSA membrane were unsialylated. The display of the data in Figure 3 allows discussion about the results obtained by MS. While characterising the two membranes, the figure highlights the differences in the amount of sialic acid. The initial investigation also revealed that the LSA membrane had a large amount of additional unsialylated polyacNAc extended structures. When investigating these unsialylated structures further we could see that it was not just the quantitative differences of sialylation between the samples but also qualitative. The addition of sialic acid typically acts as a termination event which ceases the elongation by halting further addition of lacNAc's. Hence, the presence of polyacNAc in the LSA sample may not only indicate a higher galactosyltransferase activity, but also lower sialyltransferase activity.

Discussion

We have previously described a protocol for the global enrichment of membrane proteins from mammalian cells as a method for global analysis of membrane associated oligosaccharides⁹. Here, we have used this method and LC-MS to characterise the *N*-linked oligosaccharides released from two enriched membrane fractions, which are shown to differ in their amounts of sialic acid. The LC-MS profile of both membranes were shown to differ considerably and analysis

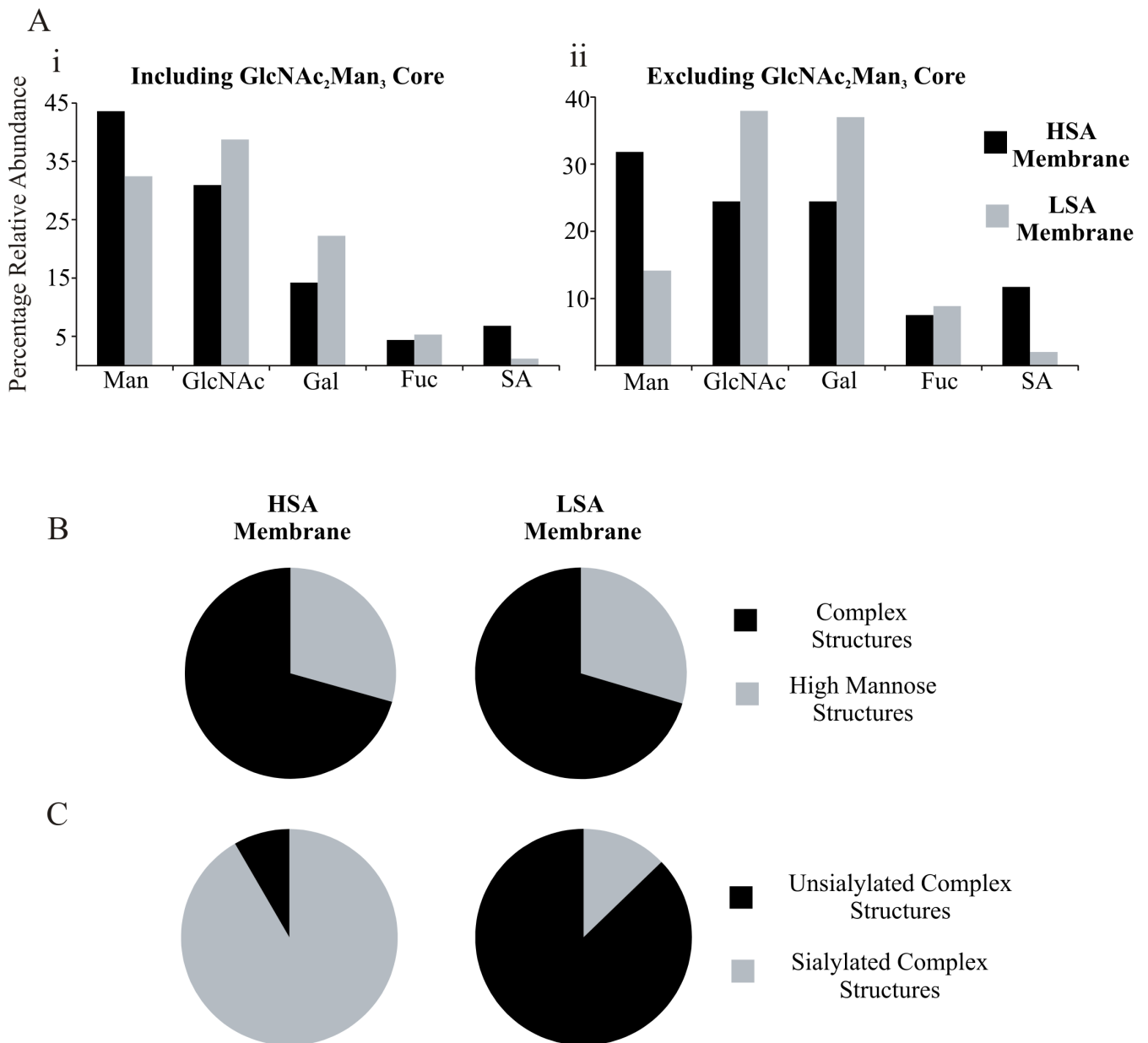


Figure 3: A) The monosaccharide composition based on the relative intensity the oligosaccharides detected on the HSA (black) and LSA (grey) membranes with the GlcNAc₂Man₃ core (i) included and (ii) excluded. Pie charts indicating the ratio of (B) high mannose and complex structures and (C) sialylated and unsialylated structures present on the HSA and LSA membranes based on relative quantification of the oligosaccharides detected.

of the released oligosaccharides revealed the presence of high-mannose and complex structures present on the cellular membranes. The presence of high-mannose structures on various cell membranes has previously been reported¹¹⁻¹⁵ and their presence here suggests that they are normal feature of membrane glycosylation. However, since plasma membranes containing the fully processed oligosaccharides are only contributing to 2-5% of the total membrane fraction¹⁶, there is the possibility of co-enrichment of premature nascent N-linked oligosaccharides that contain the N-linked intermediary and the trimmed versions of core structures present in secretory pathway in ER and Golgi or in the endocytosis and degradative pathways. This would contribute to

a portion of high-mannose structures present in both membranes samples. The role of complex oligosaccharides on membranes^{15, 17} is dependent on their structural features. Fucosylation is important for glycoprotein processing¹⁸ whereas sialylation is involved in the binding and transport of cationic molecules¹⁹ and serves as a self-recognition factors preventing activation of the immune system²⁰.

Highlighting biological importance using relative quantification. Changes in the abundance of particular oligosaccharides have been shown to have adverse effects and altered glycosylation profiles have been associated with different diseases states; for example a decrease in the abundance of high-mannose oligosaccharides on the membranes of peripheral blood mononuclear cells in patients with liver cirrhosis has been correlated with a decrease in natural killer cell activity²¹ whereas increased levels of sialylation have been associated with various tumours. The semi-quantitative data provides an overview into the glycosylation beyond merely determining the presence or absence of particular structures. In figure 2 we present the relative abundance of

the individual oligosaccharides and treat each oligosaccharide as a distinct structure. In reality, the oligosaccharides have common epitopes and thus share common biological purposes. Very rarely can a biological role be attributed to one particular oligosaccharide structure. With a lack of data representing the biological function, we have to resort to sorting the structures based on their common structural features. This approach could make more sense when communicating the glycomic results to researchers unfamiliar with MS based glycomic analysis. With the belief that structure and function are associated, sorting based on structural families is reasonable and in this case highlights the difference between samples that would require further investigation. Since interpretation of the semi-quantitative data would reveal the biological aspects of the glycosylation, we found that the type of data representation as in figure 3B and 3C as appealing, together with the MSAC metric that has been shown to be well suited for comparison of *O*-linked oligosaccharides¹⁰. With carbohydrate signalling being more analogue than digital as described above, the representation provides information of increased or decreased biological ability, rather than absence or presence of this ability

Acknowledgements

Professor Hans-Joachim Gabius, University of Munich is gratefully acknowledged for his gift cellular material. The mass spectrometers were obtained by grants from the Swedish Research Council (342-2004-4434) and from the Knut and Alice Wallenberg Foundation (KAW2007.0118).

Reference

1. Stevens, T.J. and I.T. Arkin, Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins: Structure, Function, and Bioinformatics*, 2000. 39(4): p. 417-420.
2. Gabius, H.J., Glycans: bioactive signals decoded by lectins. *Biochemical Society transactions*, 2008. 36(Pt 6): p. 1491-6.
3. Dell, A. and H.R. Morris, Glycoprotein structure determination mass spectrometry. *Science*, 2001. 291(5512): p. 2351-2356.
4. Mechref, Y. and M.V. Novotny, Structural Investigations of Glycoconjugates at High Sensitivity. *Chemical Reviews*, 2002. 102(2): p. 321-370.
5. Harvey, D.J., Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *International Journal of Mass Spectrometry*, 2003. 226(1): p. 1-35.
6. Mechref, Y. and M.V. Novotny, Miniaturized separation techniques in glycomic investigations. *Journal of Chromatography B*, 2006. 841(1-2): p. 65-78.
7. Robinson, L.J., et al., Proteomic analysis of the genetic premature aging disease Hutchinson Gilford progeria syndrome reveals differential protein expression and glycosylation. *J Proteome Res*, 2003. 2(5): p. 556-7.
8. Karlsson, N.G., et al., Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. *Rapid Commun Mass Spectrom*, 2004. 18(19): p. 2282-92.
9. Kenny, D.T., et al., Glycomic Analysis of Membrane-Associated Proteins, in *Sample Preparation in Biological Mass Spectrometry*, A.R. Ivanov and A.V. Lazarev, Editors. 2011, Springer. p. 498-513.
10. Hayes, C.A., S. Nemes, and N.G. Karlsson, Statistical analysis of glycosylation profiles to compare tissue type and inflammatory disease state. *Bioinformatics*, 2012.
11. Nuck, R., et al., Comparative study of high-mannose-type oligosaccharides in membrane glycoproteins of rat hepatocytes and different rat hepatoma cell lines. *European Journal of Biochemistry*, 1993. 216(1): p. 215-221.
12. Kim, Y.G., et al., The identification and characterization of xenotigenic nonhuman carbohydrate sequences in membrane proteins from porcine kidney. *Proteomics*, 2006. 6(4): p. 1133-42.
13. An, H.J., et al., Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of High Mannose Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells. *Molecular & Cellular Proteomics*, 2012. 11(4).
14. Penuela, S., et al., Glycosylation Regulates Pannexin Intermixing and Cellular Localization. *Molecular Biology of the Cell*, 2009. 20(20): p. 4313-4323.
15. Clark, R.A., et al., Characterisation of tissue-specific oligosaccharides from rat brain and kidney membrane preparations enriched in Na⁺,K⁺-ATPase. *Glycoconj J*, 1999. 16(8): p. 437-56.
16. Alberts, B., et al., eds. *Molecular Biology of the Cell*. 5th ed. 2008, Garland Science.
17. Li, J., et al., Processing of N-linked oligosaccharide depends on its location in the anion exchanger, AE1, membrane glycoprotein. *Biochemical Journal*, 2000. 349: p. 51-57.
18. Becker, D.J. and J.B. Lowe, Fucose: biosynthesis and biological function in mammals. *Glycobiology*, 2003. 13(7): p. 41R-53R.
19. Hu, Q.L., et al., Intracellular pathways and nuclear localization signal peptide-mediated gene transfection by cationic polymeric nanovectors. *Biomaterials*, 2012. 33(4): p. 1135-1145.
20. Pilatte, Y., J. Bignon, and C.R. Lambré, Sialic acids as important molecules in the regulation of the immune system: pathophysiological implications of sialidases in immunity. *Glycobiology*, 1993. 3(3): p. 201-218.
21. Miyaguchi, S., et al., Changes in high mannose-type glycoproteins of peripheral blood mononuclear cells in cirrhosis patients. *International Hepatology Communications*, 1995. 3(1): p. 41-

Semi-Quantitative Data Analysis for Membrane Associated *N*-linked Oligosaccharides Highlights Differences in the Glycosylation of Two Enriched Membranes Samples

Diarmuid T. Kenny^{1,2}, Kristina A. Thomsson¹, Niclas G. Karlsson¹

¹Department of Medical Biochemistry, Göteborg University, Sweden and ²School of Chemistry, National University Ireland Galway, Ireland

Glycomic data is difficult to display due to the fact that the oligosaccharides are diverse, and structural and functional aspects overlap, where an attribute cannot be explained by only one individual structure. We set out to investigate how we could transform glycomic MS data derived from the analysis of *N*-linked oligosaccharides into easily accessible information representing differences in glycosylation. In order to do this we adopted a method for the global enrichment of cellular membrane using carbonate extraction and analysed two samples with differing levels of sialylation. We could identify 41 different oligosaccharides based on their composition. By measuring the MS intensities of the compositions that were identified, we could determine the relative abundance of the individual oligosaccharides. By sorting the identified oligosaccharides by their structural characteristics (high-mannose, high and low sialylation content) and applying mass spectrometric average composition (MSAC) analysis, we could identify particular traits about the enriched membranes. Using this approach it was possible to highlight differences in glycosylation profile in order to make the data from complex MS-spectra accessible for structural comparison.

N-linked | glycosylation | LC-MS | Sialylation

The cellular membrane proteins account for up to 30 percent of the total cellular protein content and are usually found to be *N*-linked glycosylated¹. *N*-linked oligosaccharides are, for instance, present on the plasma membrane where it has been identified as a mediator of extracellular signalling and involved in immunological events². Therefore, characterising the structural features of the oligosaccharides that mediates this signalling is important. Glycomics of the immunologically important cell surface oligosaccharide epitopes can aid in our understanding of their role in immunological signalling. Mass spectrometry (MS) has revolutionised the field of glycomics and has become the primary instrument of choice for a glycomic platform³⁻⁵. When coupled to an LC-system, it is ideally suited to characterising the mixture of oligosaccharides structures typically present on a glycoprotein⁶. The data generated by glycomic investigation are usually quite complex and confusing, since similar structural features are present on several structures and the information about their biological importance is usually quite difficult to digest. The communication between the glycobiochemist and the glycoanalyst would be greatly enhanced if they conformed to using a common language and terminology of important epitopes such as high-mannose, sialylation, blood group antigens, sulfation, core types, *N*-acetylglucosamine

elongation. This would aid in the amalgamation of reports and also provide a platform for better communication regarding MS specific information that is not obtainable by other methods such as antibody binding and lectin epitopes including details about antenna distribution that serves as scaffold for the terminal epitopes and potentially semi-quantitative data relating to the abundance of oligosaccharides. In this report we investigate how to display the glycomic information obtained from MS data when comparing the glycosylation from two different mammalian membranes.

Materials and Methods

All materials were obtained from Sigma Aldrich (St Louis, MO) unless otherwise stated. The 18 MΩ water was produced using the MilliQ water purification system (Millipore, Billerica, MA)

Enrichment of Membrane proteins from mammalian cells. Membrane proteins were enriched by precipitation in sodium carbonate^{7,8}. Cellular material (100 mg) of was suspended in 100 mM sodium carbonate. The solution was sonicated on ice at 40% maximum amplitude four times for 15 seconds and cooled on ice in between each sonication for 1 minute. After sonication, the solution was transferred to a beaker and 100 mM of sodium carbonate was added to a final volume of 100 mL and stirred for 5 minutes at 4 °C. Cell debris was removed by centrifugation at 5,000 g for 10 minutes. After ultracentrifugation, the supernatant was transferred to a new centrifuge tubes and ultracentrifuged at 115,000 g for 75 minutes. The supernatant was discarded and the pellet was washed in 50 mM tris(hydroxymethyl)aminomethane-HCl (Tris-HCl) pH 7.3 and ultracentrifuged at 115,000 g for 75 minutes. The wash step was repeated once. The washed membrane containing pellets were solubilised in 7.0 M urea, 2.0 M thiourea, 40 mM Tris base and 4% w/v 3-[(3 cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS).

Release of *N*-linked oligosaccharides for LC-MS analysis. The solubilised membrane extracts were blotted onto Immobilon-P PVDF membranes (Millipore, Billerica, MA) and stained with DB71 and de-stained in 10% acetic acid in 40% ethanol. The membranes were blocked with 1% polyvinylpyrrolidone (PVP) in methanol. The *N*-linked oligosaccharides were enzymatically released from the protein by PNGase F. The enzyme (5 μL, 5 units) was added to each blot and incu-

Correspondence: Niclas G. Karlsson, Dept. Medical Biochemistry, University of Gothenburg, Box 440, 405 30 Gothenburg, SWEDEN; Phone: +46 31 786 6528. Email: Niclas.karlsson@medkem.gu.se

bated for 10 minutes at 37°C. After 10 minutes incubation a further 5 μ L of PNGase F and 10 μ L of water was added and the blots were incubated overnight at 37°C. The released oligosaccharides were reduced by the addition of 0.5 M NaBH₄ in 50 mM NaOH and incubated at 60°C overnight. The reaction was suspended with acetic acid and the samples were desalted with 60 μ L of AG 50W-X8 cation exchange beads (Bio-Rad, Hercules, CA) packed in C18 zip tips (Millipore, Billerica, MA). Borate complexes were removed by repeated addition/evaporation with 1% acetic acid in methanol (100 μ L for each addition). The released oligosaccharides were dissolved in water prior analysis by LC-MS.

LC-MS of released N-linked oligosaccharides. Released membrane associated oligosaccharides were analysed by LC-MS using a 10 cm \times 250 μ m I.D. column containing 5- μ m porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA) prepared in-house⁹. Oligosaccharides were eluted using a linear 8mM aqueous ammonium bicarbonate/acetonitrile gradient from 0-40% in 40 minutes at a flow rate of 10 μ L/minute. The eluted oligosaccharides were analysed using an ESI-IT MS (LTQ, Thermo Electron Corp., San Jose, CA) operating in negative ion mode with a spray voltage of -3.5 kV. Oligosaccharides were detected as [M-H]⁻, [M-2H]²⁻ and [M-3H]³⁻ over a scan range of *m/z* 380-2000. Individual oligosaccharides were isolated for fragmentation by collision induced dissociation (CID) with the collision energy of 35% and a dynamic integration time. The oligosaccharide composition was determined by manual interpretation of the MS and MS/MS spectral data. For relative abundance determination, compositions were identified by their *m/z* value and retention time. An extracted ion chromatogram was generated for each oligosaccharide and the MS intensity for each MS scan was summed to determine the relative abundance of the particular oligosaccharide. The MSAC analysis was performed as described by Hayes et al¹⁰. Briefly, the intensities for each composition identified were determined. The compositions were reduced to their monosaccharide content by determining the number GlcNAc, Man, Gal, Fuc and NeuAc residues. The monosaccharide content for each composition was multiplied by the relative abundance of the particular structure and then each parameter summed over the entire sample and the percentage of each summed monosaccharide was calculated to give the MSAC data. An example of this calculation is given in Supplementary Table 1.

Result

Compositional analysis of released oligosaccharides.

LC-MS has been used extensively for the analysis of oligosaccharides. Here it was utilized for the analysis of N-linked oligosaccharides released from enriched membrane proteins. Two membranes fractions were enriched from whole lysates of a high and low sialic acid (HSA and LSA) containing cells by precipitation with sodium carbonate. The LC-MS (Figure 1), shows that the glycosylation profile between the samples is very different. The first level of interpretation utilized the compositional analysis to reveal that majority of the most abundant structures in HSA contained one or two sialic acid residues, whereas the most abundant structures on the LSA were all unsialylated (Table 1) This analysis verified that the oligosaccharides differed in their amount of sialic acid. Having identified that the samples showed different profiles, the glycosylation pattern was further analysed focussing on

both high and low abundant oligosaccharides. Of the 41 identified oligosaccharide compositions, 21 were present on both membranes and 20 were detected on the LSA membranes only. Compositional analysis of these structures revealed that the oligosaccharides common to both membranes contained hexoses (Hex), N-acetylhexosamines (HexNAcs) with and without deoxyhexoses (interpreted to be fucose (Fuc)) and sialic acids (NeuAc). Additional structures present only in the LSA contained only Hex, HexNAcs with or without Fuc (see Table 1) but no sialic acid.

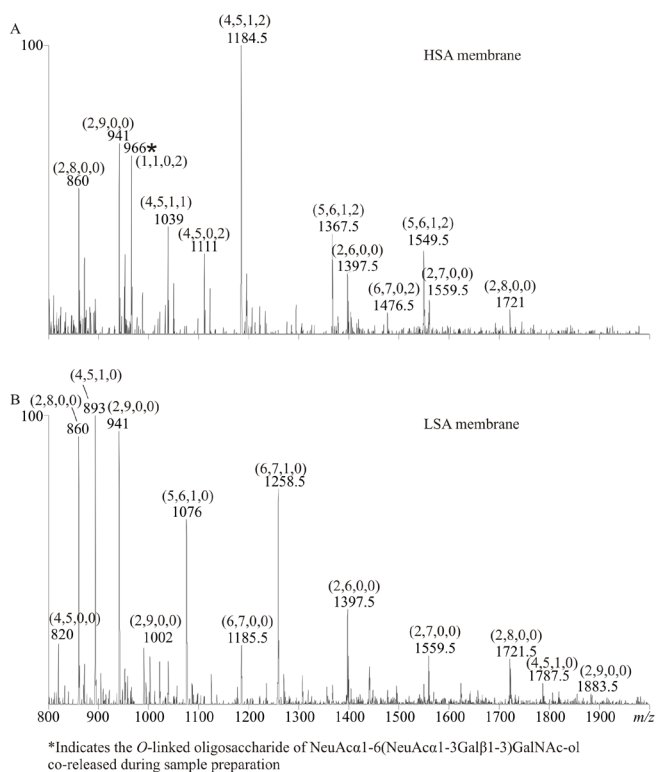


Figure 1: LC-MS of N-linked oligosaccharides released from (a) HSA and (b) LSA membranes. The oligosaccharides were detected as the [M - H]⁻, [M - 2H]²⁻, [M - 3H]³⁻ ions between 20 - 40 min. The composition of the most abundant oligosaccharides is indicated as (Hex, HexNAc, Deoxy-Hex, Sialic Acid).

Identification of high-mannose and complex structures on membranes.

The composition analysis generates a limited amount of information on the glycosylation profile of the membranes as it lacks structural details which are important when investigating the biological role of the membrane glycosylation. However, it is known that all N-linked oligosaccharides share a common GlcNAc₂Man₃ core sequence. This sequence can be elongated by the addition of more Man residues to create high-mannose structures, by the addition of N-acetylglucosamine (GlcNAc, Galβ1-4GlcNAc) residues with the possible further addition of Fuc and/or sialic acid residues to create complex structures or by a combination of both to produce hybrid structures. We can therefore utilise our knowledge of the biosynthetic pathway when assigning the structure of the oligosaccharides to garner structure specific information missing from the compositional analysis. The LC-MS data shows the presence of both high-mannose and complex structures on the membrane proteins (Table 1). The high-mannose structures were detected with 5 (Man5, *m/z* 1235), 6 (Man6, *m/z* 1397.5) and 7 (Man7, *m/z* 1559.5)

Table 1: Composition analysis of the oligosaccharides identified on the High and Low SA membrane. The composition is presented as (HexNAc, Hex, Deoxyhex, Sialic Acid). The structures for each oligosaccharides is presented using CFG annotation.

Oligosaccharides Identified on HSA and LSA				Oligosaccharides Identified on LSA only			
	[M-nH] ⁻ⁿ	Composition	Structure		[M-nH] ⁻ⁿ	Composition	Structure
1	1235 ¹⁻	(2,5,0,0)		22	911 ¹⁻	(2,3,0,0)	
2	1397.5 ¹⁻	(2,6,0,0)		23	1276 ¹⁻	(3,4,0,0)	
3	1559.5 ¹⁻	(2,7,0,0)		24	1480.5 ¹⁻	(4,4,0,0)	
4	1721.5 ¹⁻	(2,8,0,0)		25	1104 ²⁻	(6,6,0,0)	
	860 ²⁻			26	1185.5 ²⁻	(6,7,0,0)	
5	1883.5 ¹⁻	(2,9,0,0)		27	1368 ²⁻	(7,8,0,0)	
	941 ²⁻			28	1550.5 ²⁻	(8,9,0,0)	
6	1039 ²⁻	(4,5,1,1)		29	1422 ¹⁻	(3,4,1,0)	
7	1220 ²⁻	(5,5,1,1)		30	995 ²⁻	(5,5,1,0)	
8	1404 ²⁻	(6,7,1,1)		31	1177 ²⁻	(6,6,1,0)	
9	1111 ²⁻	(4,5,0,2)		32	1360 ²⁻	(8,7,1,0)	
10	1184.5 ²⁻	(4,5,1,2)		33	1441 ²⁻	(7,8,1,0)	
11	1367.5 ²⁻	(5,6,1,2)		34	1542.5 ²⁻	(8,8,1,0)	
12	1476.5 ²⁻	(6,7,0,2)		35	1623.5 ²⁻	(8,9,1,0)	
13	1549.5 ²⁻	(6,7,1,2)		36	1806.5 ²⁻	(9,10,1,0)	
14	1732.5 ²⁻	(7,8,1,2)		37	1988 ²⁻	(10,11,1,0)	
15	1915 ²⁻	(8,9,1,2)		38	1447 ³⁻	(11,12,1,0)	
16	1057 ¹⁻	(2,3,1,0)		39	1568 ³⁻	(12,13,1,0)	
17	1642.5 ¹⁻	(4,5,0,0)		40	1690.5 ³⁻	(13,14,1,0)	
	820 ²⁻			41	1812 ³⁻	(14,15,1,0)	
18	1788.5 ¹⁻	(4,5,1,0)					
	893 ²⁻						
19	1002.5 ²⁻	(5,6,0,0)					
20	1076 ²⁻	(5,6,1,0)					
21	1258.5 ²⁻	(6,7,1,0)					

Man residues as their [M-H]⁻ ions and with 8 (Man8, *m/z* 1721.5 and 860) and 9 (Man9, *m/z* 1883.5 and 941) Man residues detected as the [M-H]⁻ and/or [M-2H]²⁻ ions. The complex structures were detected as the [M-H]⁻, [M-2H]²⁻ and [M-3H]³⁻ ions (Table 1). Of the complex structures identified, 16 were detected on both membranes. These were predominantly core fucosylated with or without terminal sialylation. The additional complex structures identified on the LSA membrane were all unsialylated extended with up to 12 lacNAc residues. The data obtained by interpretation of masses into composition and biosynthetic pathways are providing the first level of interpretation and is required for stringent documentation of the raw data. This is also the minimal amount of data that is required for submission to structural databases or knowledgebases and for recording and future references.

It is then the glycoanalysts obligation to do further interpretation of the data in order to communicate the result to a wider community.

Semi-quantitative display of individual structures with high-mannose and complex type glycosylation. To better reflect the biological aspects of the membrane glycosylation, the oligosaccharides can also be displayed according to their biological characteristics. The lack of quantitative data in Table 1 is limiting the amount of biologically relevant information that can be ascertained from it. Using the predicted oligosaccharide structures from LC-MS and the intensity of each composition, it is possible generate semi-quantitative data that shows there is a significant variation in the abundance of the oligosaccharides between HSA and LSA membranes (Figure 2). Further information about the glycosylation can be obtained by sorting the oligosaccharides into high-mannose and complex type and this reveals that the main difference

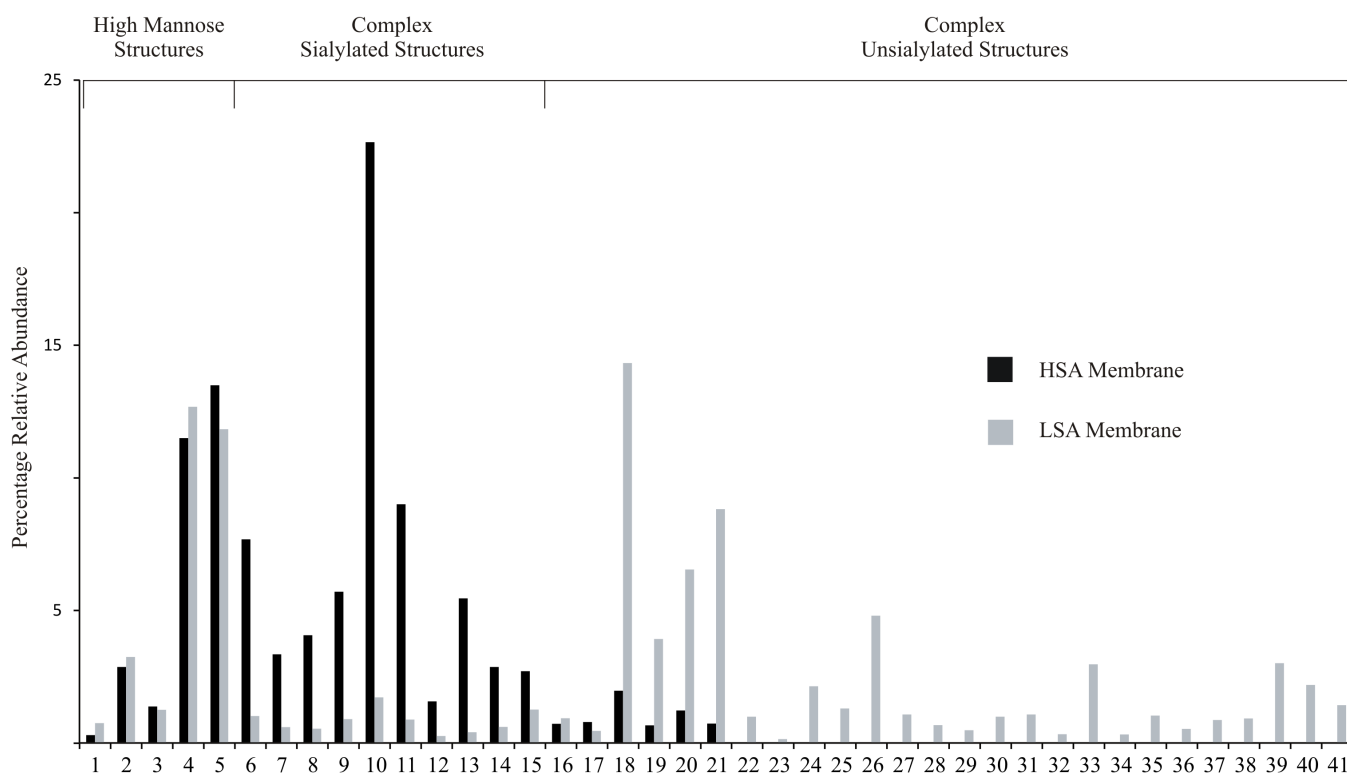


Figure 2: The relative abundance of the *N*-linked oligosaccharides detected on the HSA (back) and LSA (grey) membranes. The high mannose, complex sialylated structures and complex unsialylated structures are indicated. The number of each structure corresponds to their numbering in Table 1.

between the samples lie in the amount of sialylated and unsialylated structures. However, the MS based information is still obscuring the picture by sorting the x-axis into features that are only accessible for a trained glyco biologist, and would be difficult even for researchers in the field to digest and make the connection between the identified structures and their function.

Average based MS data to display glycosylation differences. We have described a method for how to use the structural information from MS and the intensities to calculate a mass spectrometry average composition (MSAC)¹⁰ (Figure 3A). The monosaccharide content of the membranes, based on the semi-quantitative data, can be used to display the glycosylation within a global context. This approach used for displaying *O*-linked semi-quantitative data was found to be superior in highlighting the difference in the amount of *N*-linked sialic acid present on both membranes. Furthermore, from the data we can see a higher amount of HexNAc and Hex in the complex oligosaccharides residues, indicating a higher abundance of lacNAc residues. For *O*-linked data there are a number of cores that could be terminated or extended, but *N*-linked oligosaccharides all have a common core. In figure 3Ai the monosaccharide content was determined with the inclusion of the GlcNAc₂Man₃ core sequence. However, since all structures contain this core structures, the MS based monosaccharide content data was normalized by excluding the core sequence to stress the differences in the antennas (Fig 3Aii). Combining this MSAC method with info about *N*-linked glycosylation (complex, high-mannose) provides an efficient way to grasp the MS data. In figure 3B we

can see that the abundance of high-mannose and complex structures for both membranes is similar and that the majority of the oligosaccharides in both samples are complex structures. Of those complex structures, the sialylated structures are the major component on the HSA membrane, whereas they are only a minor component on the LSA membrane (Figure 3C). The relative abundance of the individual high-mannose oligosaccharides did not vary greatly between the samples, the complex oligosaccharides differed significantly with the sialylated structures having greater abundance on HSA membranes than LSA membranes. Conversely the most abundance structures on the LSA membrane were unsialylated. The display of the data in Figure 3 allows discussion about the results obtained by MS. While characterising the two membranes, the figure highlights the differences in the amount of sialic acid. The initial investigation also revealed that the LSA membrane had a large amount of additional unsialylated polyacNAc extended structures. When investigating these unsialylated structures further we could see that it was not just the quantitative differences of sialylation between the samples but also qualitative. The addition of sialic acid typically acts as a termination event which ceases the elongation by halting further addition of lacNAc's. Hence, the presence of polyacNAc in the LSA sample may not only indicate a higher galactosyltransferase activity by promoting the elongation, but also lower sialyltransferase activity.

Discussion

We have previously described a protocol for the global enrichment of membrane proteins from mammalian cells as a method for global analysis of membrane associated oligosaccharides⁹. Here, we have used this method and LC-MS to characterise the *N*-linked oligosaccharides released from two enriched membrane fractions, which are shown to differ in their amounts of sialic acid. The LC-MS profile of both membranes were shown to differ considerably and analysis

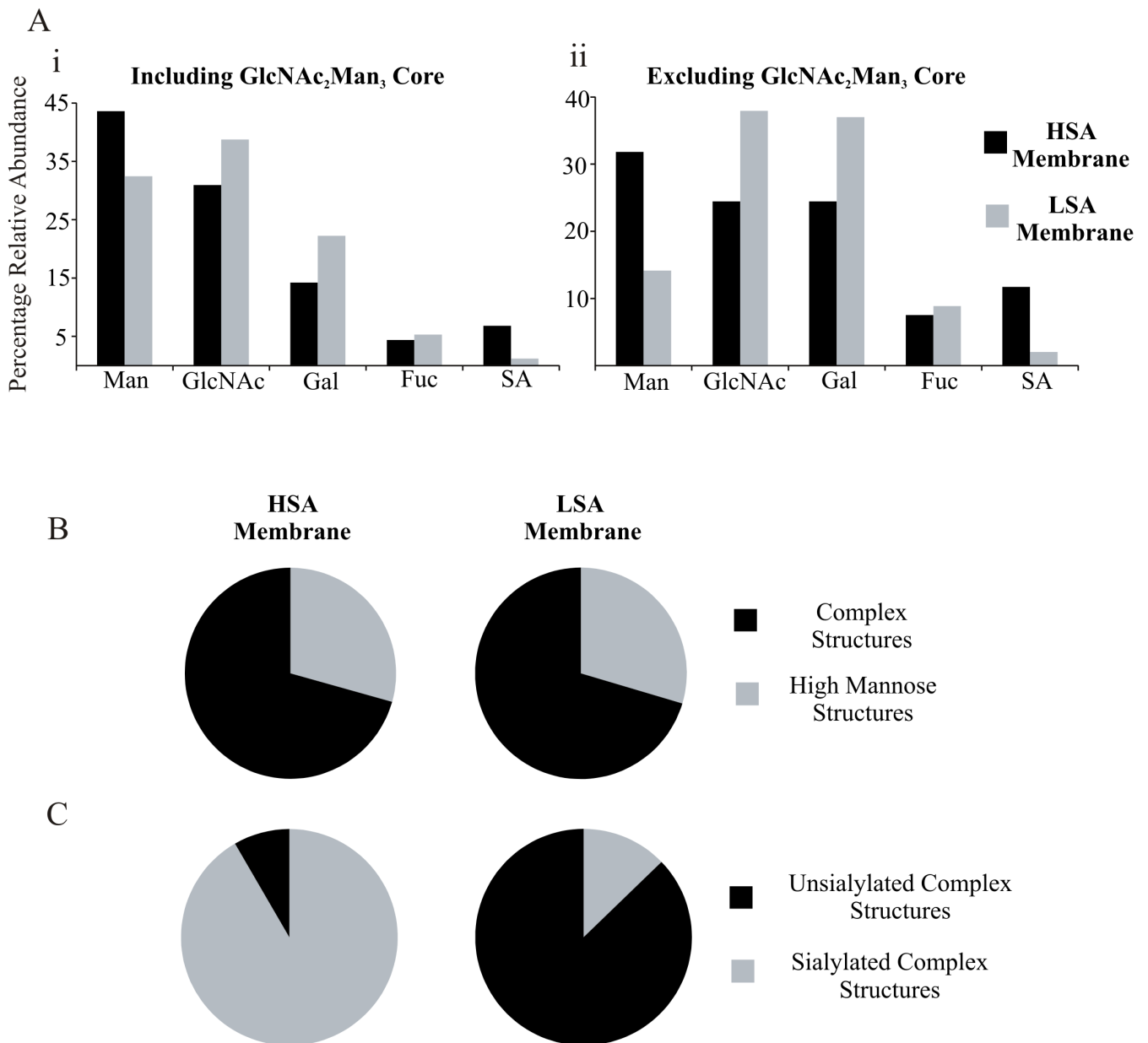


Figure 3: A) The monosaccharide composition based on the relative intensity the oligosaccharides detected on the HSA (black) and LSA (grey) membranes with the GlcNAc₂Man₃ core (i) included and (ii) excluded. Pie charts indicating the ratio of (B) high mannose and complex structures and (C) sialylated and unsialylated structures present on the HSA and LSA membranes based on relative quantification of the oligosaccharides detected.

of the released oligosaccharides revealed the presence of high-mannose and complex structures present on the cellular membranes. The presence of high-mannose structures on various cell membranes has previously been reported¹¹⁻¹⁵ and their presence here suggests that they are normal feature of membrane glycosylation. However, since plasma membranes containing the fully processed oligosaccharides are only contributing to 2-5% of the total membrane fraction¹⁶, there is the possibility of co-enrichment of premature nascent N-linked oligosaccharides that contain the N-linked intermediary and the trimmed versions of core structures present in secretory pathway in ER and Golgi or in the endocytosis and degradative pathways. This would contribute to

a portion of high-mannose structures present in both membranes samples. The role of complex oligosaccharides on membranes^{15, 17} is dependent on their structural features. Fucosylation is important for glycoprotein processing¹⁸ whereas sialylation is involved in the binding and transport of cationic molecules¹⁹ and serves as a self-recognition factors preventing activation of the immune system²⁰.

Highlighting biological importance using relative quantification. Changes in the abundance of particular oligosaccharides have been shown to have adverse effects and altered glycosylation profiles have been associated with different diseases states; for example a decrease in the abundance of high-mannose oligosaccharides on the membranes of peripheral blood mononuclear cells in patients with liver cirrhosis has been correlated with a decrease in natural killer cell activity²¹ whereas increased levels of sialylation have been associated with various tumours. The semi-quantitative data provides an overview into the glycosylation beyond merely determining the presence or absence of particular structures. In figure 2 we present the relative abundance of

the individual oligosaccharides and treat each oligosaccharide as a distinct structure. In reality, the oligosaccharides have common epitopes and thus share common biological purposes. Very rarely can a biological role be attributed to one particular oligosaccharide structure. With a lack of data representing the biological function, we have to resort to sorting the structures based on their common structural features. This approach could make more sense when communicating the glycomic results to researchers unfamiliar with MS based glycomic analysis. With the belief that structure and function are associated, sorting based on structural families is reasonable and in this case highlights the difference between samples that would require further investigation. Since interpretation of the semi-quantitative data would reveal the biological aspects of the glycosylation, we found that the type of data representation as in figure 3B and 3C as appealing, together with the MSAC metric that has been shown to be well suited for comparison of *O*-linked oligosaccharides¹⁰. With carbohydrate signalling being more analogue than digital as described above, the representation provides information of increased or decreased biological ability, rather than absence or presence of this ability

Acknowledgements

Professor Hans-Joachim Gabius, University of Munich is gratefully acknowledged for his gift cellular material. The mass spectrometers were obtained by grants from the Swedish Research Council (342-2004-4434) and from the Knut and Alice Wallenberg Foundation (KAW2007.0118).

Reference

1. Stevens, T.J. and I.T. Arkin, Do more complex organisms have a greater proportion of membrane proteins in their genomes? *Proteins: Structure, Function, and Bioinformatics*, 2000. 39(4): p. 417-420.
2. Gabius, H.J., Glycans: bioactive signals decoded by lectins. *Biochemical Society transactions*, 2008. 36(Pt 6): p. 1491-6.
3. Dell, A. and H.R. Morris, Glycoprotein structure determination mass spectrometry. *Science*, 2001. 291(5512): p. 2351-2356.
4. Mechref, Y. and M.V. Novotny, Structural Investigations of Glycoconjugates at High Sensitivity. *Chemical Reviews*, 2002. 102(2): p. 321-370.
5. Harvey, D.J., Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *International Journal of Mass Spectrometry*, 2003. 226(1): p. 1-35.
6. Mechref, Y. and M.V. Novotny, Miniaturized separation techniques in glycomic investigations. *Journal of Chromatography B*, 2006. 841(1-2): p. 65-78.
7. Robinson, L.J., et al., Proteomic analysis of the genetic premature aging disease Hutchinson Gilford progeria syndrome reveals differential protein expression and glycosylation. *J Proteome Res*, 2003. 2(5): p. 556-7.
8. Karlsson, N.G., et al., Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. *Rapid Commun Mass Spectrom*, 2004. 18(19): p. 2282-92.
9. Kenny, D.T., et al., Glycomic Analysis of Membrane-Associated Proteins, in *Sample Preparation in Biological Mass Spectrometry*, A.R. Ivanov and A.V. Lazarev, Editors. 2011, Springer. p. 498-513.
10. Hayes, C.A., S. Nemes, and N.G. Karlsson, Statistical analysis of glycosylation profiles to compare tissue type and inflammatory disease state. *Bioinformatics*, 2012.
11. Nuck, R., et al., Comparative study of high-mannose-type oligosaccharides in membrane glycoproteins of rat hepatocytes and different rat hepatoma cell lines. *European Journal of Biochemistry*, 1993. 216(1): p. 215-221.
12. Kim, Y.G., et al., The identification and characterization of xenotigenic nonhuman carbohydrate sequences in membrane proteins from porcine kidney. *Proteomics*, 2006. 6(4): p. 1133-42.
13. An, H.J., et al., Extensive Determination of Glycan Heterogeneity Reveals an Unusual Abundance of High Mannose Glycans in Enriched Plasma Membranes of Human Embryonic Stem Cells. *Molecular & Cellular Proteomics*, 2012. 11(4).
14. Penuela, S., et al., Glycosylation Regulates Pannexin Intermixing and Cellular Localization. *Molecular Biology of the Cell*, 2009. 20(20): p. 4313-4323.
15. Clark, R.A., et al., Characterisation of tissue-specific oligosaccharides from rat brain and kidney membrane preparations enriched in Na⁺,K⁺-ATPase. *Glycoconj J*, 1999. 16(8): p. 437-56.
16. Alberts, B., et al., eds. *Molecular Biology of the Cell*. 5th ed. 2008, Garland Science.
17. Li, J., et al., Processing of N-linked oligosaccharide depends on its location in the anion exchanger, AE1, membrane glycoprotein. *Biochemical Journal*, 2000. 349: p. 51-57.
18. Becker, D.J. and J.B. Lowe, Fucose: biosynthesis and biological function in mammals. *Glycobiology*, 2003. 13(7): p. 41R-53R.
19. Hu, Q.L., et al., Intracellular pathways and nuclear localization signal peptide-mediated gene transfection by cationic polymeric nanovectors. *Biomaterials*, 2012. 33(4): p. 1135-1145.
20. Pilatte, Y., J. Bignon, and C.R. Lambré, Sialic acids as important molecules in the regulation of the immune system: pathophysiological implications of sialidases in immunity. *Glycobiology*, 1993. 3(3): p. 201-218.
21. Miyaguchi, S., et al., Changes in high mannose-type glycoproteins of peripheral blood mononuclear cells in cirrhosis patients. *International Hepatology Communications*, 1995. 3(1): p. 41-

Supplementary Data 1: Determination of MSAC for HSA from the cumulative intensities of all the structures detected. The MSAC analysis reduces the samples to their monosaccharide content.

<i>m/z</i> value of the structures detected on HSA <i>m/z</i>	Combined intensities of all structures with the same <i>m/z</i> value Signal Intensity	Number of residues per <i>m/z</i> value					The intensity of each monosaccharide residues per <i>m/z</i> value (intensity X number of residues)				
		Man	Gal	glcNAc	Fuc	SA	Man	Gal	glcNAc	Fuc	SA
1235	605.2	5	0	2	0	0	263213	0	105285.2	0	0
1397	6527.4	6	0	2	0	0	302587	0	100862.2	0	0
1559	3092.5	7	0	2	0	0	123169	0	35191.2	0	0
1721	52642.6	8	0	2	0	0	288312	0	72078	0	0
1883	50431.1	9	0	2	0	0	68634	0	15252	0	0
1038	17595.60	3	2	4	1	1	61893.6	41262.4	82524.8	20631.2	20631.2
1183	36039.00	3	2	4	1	2	30132	20088	40176	10044	20088
1220	7626.00	3	3	5	1	1	72843	72843	121405	24281	24281
1367	20631.20	3	3	5	1	2	10602	10602	17670	3534	7068
1111	10044.00	3	2	4	0	2	37395	24930	49860	0	24930
1404	24281.00	3	6	4	1	1	4750.8	9501.6	6334.4	1583.6	1583.6
1476	3534.00	3	4	6	0	2	13365.3	17820.4	26730.6	0	8910.2
1549	12465.00	3	4	6	1	2	8265.6	11020.8	16531.2	2755.2	5510.4
1057	1583.60	3	0	4	1	0	4843.8	0	6458.4	1614.6	0
893	4455.10	3	2	4	1	0	13365.3	8910.2	17820.4	4455.1	0
1075	2755.20	3	3	5	1	0	8265.6	8265.6	13776	2755.2	0
1258	1614.60	3	4	6	1	0	4843.8	6458.4	9687.6	1614.6	0
820	1737	3	2	4	0	0	5211	3474	6948	0	0
1002	1458.2	3	5	3	0	0	4374.6	7291	4374.6	0	0
1732	6512	3	5	7	1	0	19536	32560	45584	6512	0
							1.4E+06	4.9E+05	9.2E+05	1.4E+05	2.1E+05

Cumulative intensities for each monosaccharide residue

Presence of terminal *N*-acetylgalactosamine β 1-4*N*-acetylglucosamine residues on *O*-linked oligosaccharides from gastric MUC5AC: Involvement in *Helicobacter pylori* colonization?

Diarmuid T Kenny^{2,3}, Emma C Skoog^{3,4}, Sara K Lindén^{3,4}, Weston B Struwe⁵, Pauline M Rudd⁵, and Niclas G Karlsson^{1,3}

²School of Chemistry, National University of Ireland, Galway, Ireland; ³Medical Biochemistry, University of Gothenburg, PO Box 440, 405 30 Gothenburg, Sweden; ⁴Mucosal Immunobiology and Vaccine Center, Sahlgrenska Academy, University of Gothenburg, 405 30 Gothenburg, Sweden; and ⁵National Institute for Bioprocessing Research and Training, University College Dublin, Dublin, Ireland

Received on October 21, 2011; revised on March 5, 2012; accepted on April 21, 2012

Isolation of MUC5AC mucins from the gastric mucosa from two secretor individuals (one from normal mucosa from a patient with gastric cancer and one from a control) showed different abilities to bind and induce the proliferation of the *Helicobacter pylori* strain J99. Analysis of the released *O*-linked oligosaccharides by LC-MS from these individuals showed a very heterogeneous mixture of species from the cancer patient containing both neutral and sialylated structures, whereas the normal sample showed dominating neutral blood group H terminating structures as well as neutral structures containing the di-*N*-acetylglucosamine (lacdiNAc) unit GalNAc β 1-4GlcNAc β 1- on the C-6 branch of the reducing end GalNAc. The linkage configuration of these epitopes were determined using C-4-specific fragmentation for the GalNAc β 1-4GlcNAc β 1- glycosidic linkage, comparison of the MS³ fragmentation with standards for linkage configuration and *N*-acetylhexosamine type as well as exoglycosidase treatment. It was also shown that the lacdiNAc epitope is present in both human and porcine gastric mucins, indicating that this is an epitope preserved between species. We hypothesize that the termination on gastric MUC5AC with lacdiNAc is in competition with complex glycosylation such as the Le^b and H type 1 as well as complex sialylated structures. These are epitopes known to bind the *H. pylori* BabA and SabA adhesins.

Keywords: BabA / gastric cancer / *Helicobacter* / lacdiNAc / mucin / SabA

Introduction

The gastro intestinal tract is covered in a semi-permeable mucus layer that primarily consists of secreted mucins. This layer serves to protect the gastric epithelial cells from potential damage that could occur due to mechanical and chemical stress or enzymatic activity within the tract. The mucus layer is composed of oligomeric mucin glycoproteins that serve as a scaffold for an abundance of anti-microbial epitopes (McGuckin et al. 2011). The mucin layer can also act as a functional binding site for certain pathogens such as *H. pylori*. *H. pylori* are a spiral-shaped gram-negative bacteria that colonizes the mucus layer covering the gastric epithelium. The bacteria were first isolated and cultured from the gastric mucosa by Marshall and Warren in 1983 and their discovery revealed the *H. pylori*'s association with gastritis, peptic ulcers and several forms of gastric cancer (Marshall and Warren 1984). *H. pylori* infection is the most common bacterial infection worldwide and it is estimated to affect over 50% of the world's population (Kobayashi et al. 2009). Infection usually begins in early childhood and infection and its associated symptoms such as chronic gastritis and peptic ulcers can persist throughout an individual's life. *H. pylori* infection can dramatically increase the risk of gastric cancer and has been attributed as a key mitigating factor for up to 65% of gastric cancers, equivalent to 5.5% of all cancers globally (Menaker et al. 2004).

Gastric mucins contain highly diverse carbohydrates that can be utilized by *H. pylori* for adhesion. The attachment of *H. pylori* is mediated by outer membrane proteins (OMPs) with lectin-like properties. The adhesion of *H. pylori* to the gastric surface is mediated by both the presence of these adhesion OMPs on the surface of the *H. pylori* cell and the presence of particular glyco-epitopes on the mucin. Of the OMPs that are involved in attachment, two have been identified as particularly important for the attachment of *H. pylori* to the mucosal layer: blood group antigen-binding adhesin (BabA) and the sialic acid-binding adhesin (SabA). BabA recognizes

¹To whom correspondence should be addressed: Tel: +46-31-786-6528; Fax: +46-31-41-6108; e-mail: niclas.karlsson@medkem.gu.se

the Le^b and H type-1 epitopes (Ilver et al. 1998) and SabA recognizes sialylated oligosaccharides (Mahdavi et al. 2002).

In the adult stomach, the primary secreted mucins on the gastric wall are MUC5AC and MUC6, which harbor a vast number of different glyco-structures which are predominantly neutral in healthy tissue (Nordman et al. 2002; McGuckin et al. 2011). Therefore, BabA-mediated binding to MUC5AC via Le^b containing mucin structures is the dominating mode of adhesion (Lindén et al. 2002; Styer et al. 2010). During the course of *H. pylori* infection, inflammation and cancer development, the mucin glycosylation can change, displaying more sialylated and sulfated structures on the mucins (Sakamoto et al. 1989; Cooke et al. 2009). Binding through SabA, therefore, also occur after colonization and subsequent inflammation (Mahdavi et al. 2002). The glands of the gastric mucosa have also been shown to express the oligosaccharide epitope GlcNAc α 1-4Gal. The presence of this epitope has been indicated to suppress the growth of *H. pylori* (Kawakubo et al. 2004). In contrast, binding to MUC5AC increase *H. pylori* proliferation (Skoog et al. 2011).

We have previously used MUC5AC and MUC6 isolated from the stomachs with and without cancer to investigate the relationship between the mucin glycosylation and *H. pylori* binding and proliferation (Skoog et al. 2011). Binding assays using a wild-type J99 strain of *H. pylori* was used to assess BabA mediated for binding to mucins from the healthy and cancerous stomachs. From this sample set, we identified one interesting characteristic secretor sample that neither bound *H. pylori* nor promoted the proliferation of the bacteria. We were interested in identifying key glyco-epitopes that were present on this healthy tissue differentiating it from the normal tissue from patients with gastric cancer. By identifying key glycosylation characteristics, it may be possible to use this information to predict an individual susceptibility to *H. pylori* infection and thus develop preventative measures to avoid chronic infection.

Results

We compared the glycosylation of MUC5AC from two secretor individuals, which interacted differently with *H. pylori*.

One human gastric mucin (HGM) sample (HGM-1, MUC5AC isolated from the normal mucosa of a patient with gastric cancer) was found to bind to, and promote, the proliferation of the *H. pylori* strain J99 in a BabA/Le^b-dependent manner, whereas the other sample (HGM-2, MUC5AC isolated from a healthy individual) was negative for both characteristics and did not have any apparent Le^b activity, despite expressing blood group H (Figure 1 and Tables I and II). We aimed to identify particular features of the glycosylation of HGM-2 that could explain this negative behavior.

LC-MS analysis of mucin oligosaccharides

To look at the structural differences between both samples, the *O*-linked oligosaccharides from MUC5AC isolated from stomach biopsies were released by reductive β -elimination analyzed by LC-MS. The two MS spectra were shown to be very different (Figure 2). The MS spectrum showed that the glycosylation of HGM-1 (Figure 2A) was very heterogeneous, dominated by neutral and sialylated structures containing both Lewis-type fucosylation and blood group A, B and H epitopes (Supplementary data, Table SI). This is contrasted to the HGM-2 glycosylation (Figure 2B) that was dominated by a small number of neutral structures and low levels of sialylation. The LC-MS² analysis of this samples verified that this was indeed a secretor individual, i.e. the [M-H]⁻ ions of *m/z* 530, 733 and 1041 gave LC-MS² spectra identical to the MS² spectra of blood group H containing structures reported in the MS² database UniCarb-DB (Hayes et al. 2011; Table II). The spectra showed confirmative A-fragment ions for blood group H type 2 and H type 3 (Karlsson et al. 2004) and LC-MS retention times (RTs) consistent with the biological standards of the assigned blood group H structures. In addition to these structure, we also found a prominent signal from an [M-H]⁻ ion of *m/z* 936 corresponding to a composition of three HexNAc residues, one Hex residue and one Fuc residue. In HGM-2, this composition was mainly made up of one dominating isomer as shown by the selected ion chromatogram (SIC) showing a peak at RT 24.1 min corresponding to 95% of the signal from the *m/z* 936 positive peaks throughout the SIC. This isomer was also present in the HGM-1 sample but it only corresponded to 30% of the isomers. With a

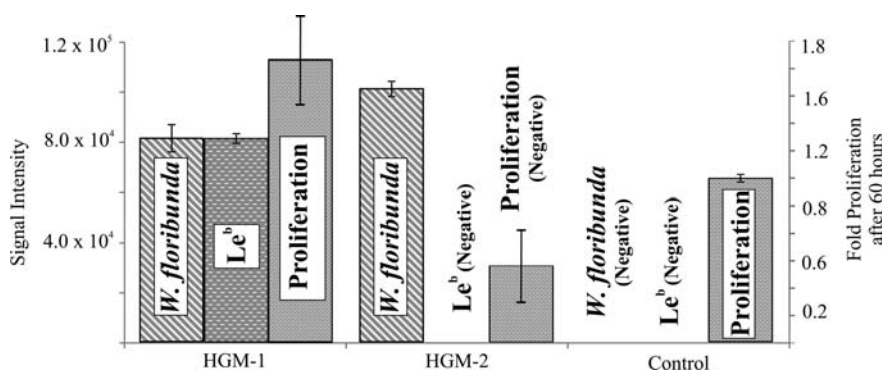


Fig. 1. Immunodetection of Le^b and terminal GalNAc on gastric MUC5AC. MUC5AC containing HGM-1 and HGM-2 was blotted onto PVDF membranes and probed with biotinylated lectin from *W. floribunda* and antibody against Le^b. Bovine fetuin used as a negative control showed no staining with either of the probes. Table I shows the characteristics of the mucins and their effect on *H. pylori* strain J99 as published (Skoog et al. 2011).

Table I. Characteristics of isolated HGMS

Sample name	Tissue status	Mucin type		Le ^b	Blood group	<i>Helicobacter pylori</i>	
		MUC5AC	MUC6			Prolif.	Binding
HGM-1	Normal tissue, cancer patient	+	Trace	+	AB	+	+
HGM-2	Normal tissue, healthy patient	+	Trace	-	O,H	-	-

differential expression of these isomers between the samples and with the high overall abundance of the RT 24.1 min isomer in the HGM-2 sample that did not bind to *H. pylori* J99 or promote its proliferation, we decided to further investigate the *m/z* 936 isomer corresponding to this composition, considering that it was unusually high in HexNAc residues compared with most *O*-linked oligosaccharides.

Identification of the di-N-acetyllactosamine epitope on human gastric MUC5AC

In order to further investigate the features of the HGM-2 glycosylation and in particular the abundant isomer found to be differentially expressed, LC-MS² was performed and the sequences interpreted. The LC-MS² spectrum of the RT 24.1 min *m/z* 936 isomer is shown in Figure 3A and the MS² of this structure showed the dominating Z fragment of *m/z* 610 and a lower intense ion corresponding to the Y fragment of *m/z* 628. An intense *m/z* 610 fragment ion in the MS² spectrum was also detected for the less intense [M-H]⁻ ion of *m/z* 790 in the MS spectrum of oligosaccharides released from HGM-2 (data not shown). It has previously been shown that the loss of the C-3 extension of the reducing end GalNAcol provides the most prominent fragmentation ion (Karlsson et al. 2004). In the case of the *m/z* 936 isomer shown in Figure 3, this loss to the fragment ion *m/z* 610 corresponds to the mass of a Fuc (*m/z* 146) and a Hex (*m/z* 162) residue. The low intense Y fragment ion of *m/z* 790 in Figure 3A showed that the C-3 branch of the GalNAcol was arranged in a sequence corresponding to the blood group H sequence Fuc α 1-2Gal β 1-. The remaining oligosaccharide after the loss of blood group H type provided a composition corresponding to two HexNAcs and the reducing end GalNAcol. In conclusion, the analysis of the MS² spectrum indicated that the overall structure has a core 2 type sequence Gal β 1-3(GlcNAc β 1-6)GalNAc that is terminating on the C-3 branch with a Fuc residue giving a blood group H type 3 sequence (Fuc α 1-2Gal β 1-3GalNAc) and the C-6 branch terminating with the addition of a HexNAc residue on the core 2 GlcNAc (Table II). Further insight into the nature of this terminating residue was provided by cross ring ^{0,2}A fragments of the core 2 GlcNAc residue, where extension to the C-4 provides a diagnostic ion of *m/z* 304 after loss of water, whereas the extension of C-3 does not give this fragment (Karlsson et al. 2004). This shows that the structure is Fuc α 1-2Gal β 1-3(HexNAc1-4GlcNA β 1-6)GalNAcol. The 1-4 linkage was also

Table II. Structures identified on the HGM-2 samples neither binding to nor inducing the proliferation of the *H. pylori* strain J99

[M-H] ⁻	Structures ^a
530 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3
733 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-4GlcNAc β 1-3 GlcNAc β 1-6
733 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3 GalNAcβ1-4GlcNAcβ1-6
790 ⁻	GalNAc-ol Gal β 1-3 Gal β 1-4/3(Fuc α 1-3/4)GlcNAc β 1-6
895 ⁻	GalNAc-ol Gal β 1-3 Fuc α 1-2Gal β 1-4GlcNAc β 1-6
895 ⁻	GalNAc-ol Gal β 1-3 Gal β 1-4GlcNAc β 1-6
895 ⁻	GalNAc-ol Gal β 1-3 Fuc α 1-2Gal β 1-3 GalNAcβ1-4GlcNAcβ1-6
936 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3 Fuc α 1-2Gal β 1-4GlcNAc β 1-6
1041 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3 Fuc α 1-2Gal β 1-4GlcNAc β 1-6
1186 ⁻	GalNAc-ol NeuAc α 2Gal β 1-3 NeuAc α 2-Gal β 1-4/3GlcNAc β 1-6
1186 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3 Gal β 1-4GlcNAc β 1-6
1406 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3/4GlcNAc β 1-3(Fuc α 1-2)Gal β 1-3 Fuc α 1-2Gal β 1-4GlcNAc β 1-6
1406 ⁻	GalNAc-ol Gal β 1-3/4GlcNAc β 1-3(Fuc α 1-2)Gal β 1-3 Fuc α 1-2Gal β 1-4GlcNAc β 1-6
1552 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-4GlcNAc β 1-3(Fuc α 1-2)Gal β 1-3 Fuc α 1-2Gal β 1-4GlcNAc β 1-3(Fuc α 1-2)Gal β 1-4GlcNAc β 1-6
1552 ⁻	GalNAc-ol Fuc α 1-2Gal β 1-3

^aSequences identified by comparison of MSⁿ spectra from UniCarb-DB (www.unicarb-db.org). Linkage configuration and positions are based on core 1, 2 and 3 series of extension. Type 2 configuration Gal β 1-4GlcNAc β 1-3 is based on the detection of C-4 linkage-specific ^{0,2}A fragment ions. Terminal epitope GlcNAc β 1-4GlcNAc is assumed based on evidence in this report and the presence of the ^{0,2}A fragment ion of *m/z* 304 in HexNAc-HexNAc sequences confirming the linkage position. Fuc configuration is assumed to be α 1-2 if present in a blood group H sequence and α 1-3/4 if linked to GlcNAc. GalNAc α 1-3 and Gal α 1-3 is assumed in blood group AB sequences. Hex always assumed to be Gal and HexNAc to be GlcNAc except in blood group A. Reducing end HexNAc-ol assumed to be GalNAc-ol. Bolded and underlined sequences correspond to C-6 terminating lacdiNac.

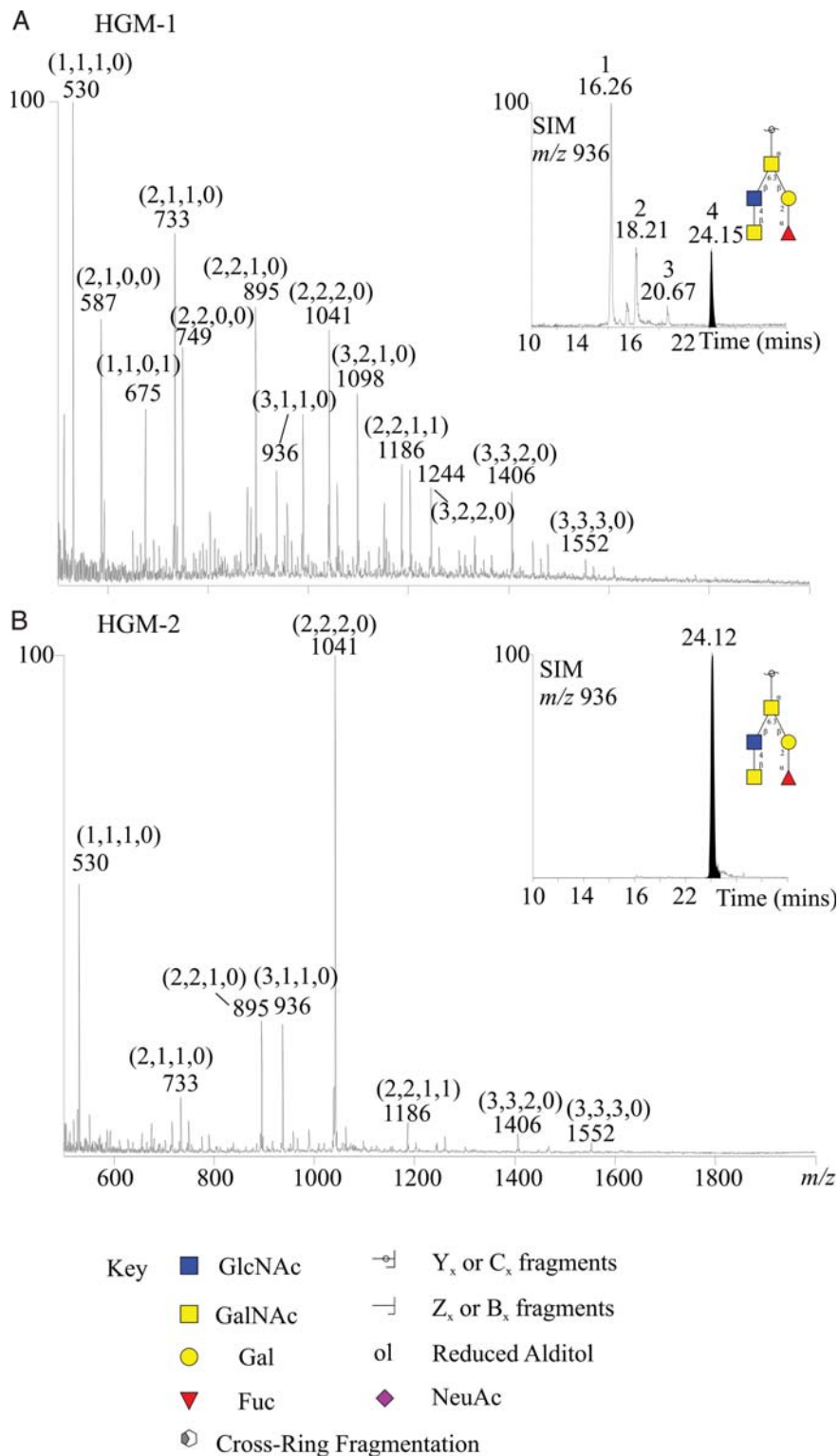


Fig. 2. LC-MS of oligosaccharides released from human gastric Muc5AC. Oligosaccharides detected as $[M-H]^-$ ions between 10 and 30 min from HGM-1 (non-tumor tissue from gastric cancer patient) (A) and HGM-2 (normal tissue) (B), from two secretor individuals. Oligosaccharide composition is labeled as [HexNAc, Hex, Fuc, NeuAc]. Inserted is the SIC of m/z 936, where the lactiNAc containing component Fuc α 1-2Gal β 1-3(GalNAc β 1-4GlcNAc β 1-6)GalNAcol eluting at 24.1 min is indicated (black). The confirmation of the lactiNAc epitope as described in the text and in Figures 3 and 4. Other isomers detected in the SIC are labeled with RT. These peaks are also numbered, with the number corresponding to their respective structures from Supplementary data, Table S1. Unlabelled peaks correspond to signals from non-isomeric oligosaccharides.

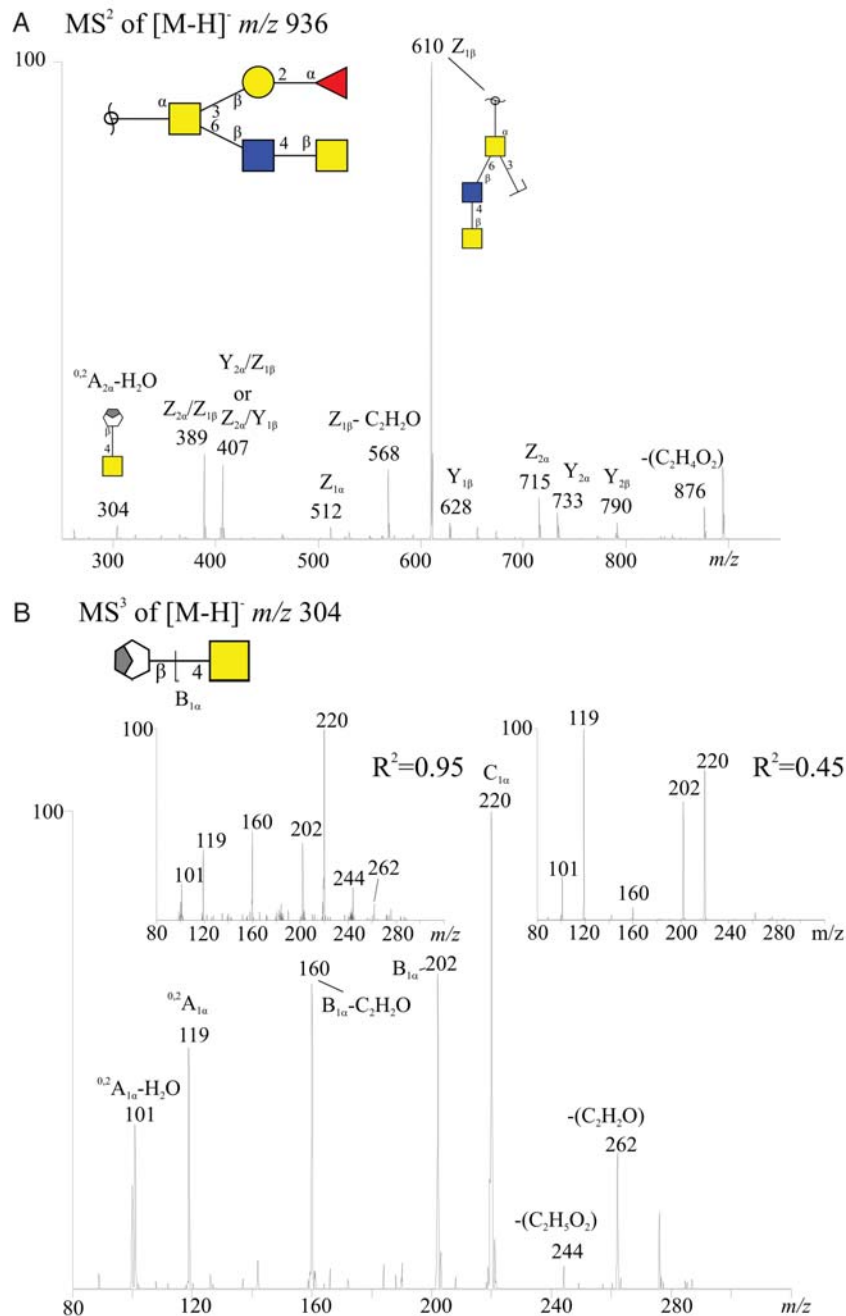


Fig. 3. Identification of the lacdiNAc sequence by MSⁿ. The LC-MS² fragmentation of the component Fuc α 1-2Gal β 1-3(GalNAc β 1-4GlcNAc β 1-6)GalNAcol in Figure 1 to identify the C-6 branch of the GalNAcol ($Z_{1\beta}$ fragment ion of *m/z* 610) and the linkage position of the GalNAc β 1-4GlcNAc β 1- moiety ($^{0,2}A_{1\alpha}$ -H₂O C-4-specific fragment ion of *m/z* 304) (A). (B) The MS³ fragmentation of the $^{0,2}A_{1\alpha}$ -H₂O ion is seen containing the GalNAc β 1-4 moiety plus part of the cleaved GlcNAc. Inserts show the MS³ fragmentation of the $^{0,2}A_{1\alpha}$ -H₂O fragment ion of *m/z* 304 isolated from GalNAc β 1-4Gal (left) and GlcNAc β 1-4GlcNAc β 1-4GlcNAc, showing with the *R*² values that the linkage of configuration of the sample correspond to the GalNAc1-4 linkage of the standard. Confirmation of the β -linkage of the lacdiNAc moiety as described in Figure 4. Key for symbols, see Figure 2.

confirmed using MS² fragmentation after permethylation (data not shown).

Having identified that HGM can be terminated by a di-HexNAc epitope, we set out to uncover the identity of the terminal 4-linked HexNAc of the structure. Further investigations, using both structural databases and literature, showed that

certain structural motifs with two HexNAcs have been identified in the gastric system. These included the GlcNAc β 1-4GlcNAc chitobiose (Newman and Kabat 1976) and the GalNAc β 1-4GlcNAc di-*N*-acetylglucosamine (lacdiNAc) motifs (Ikehara et al. 2006). An MS³ approach was adopted, whereby the fragmentation pattern of the known standards containing

GlcNAc β 1-4 and GalNAc β 1-4 was compared with the fragmentation of the di-HexNAc structure identified in the dominating m/z 936 isomer. The cross ring $^{0,2}A$ fragment with an m/z 304 characteristic for the C-4 extension of the core 2 GlcNAc (Figure 3B) was selected for MS³ fragmentation for both the sample and the standards. This particular fragment retained the full terminal HexNAc as well as the linkage position to the second outermost monosaccharide residue on the reducing end. Figure 3B shows the spectra for the MS³ of HGM-2 with the m/z 936 parent and also the subsequent collision of the daughter ion m/z 304 and the MS³ spectra of the daughter ion m/z 304 after the collision of the [M-H]⁻ ion for both standards. Correlation of the MS³ fragments and their intensities from the gastric sample with the standard showed that the standard oligosaccharide with the β 1-4 GlcNAc had an R^2 value of 0.49, whereas the β 1-4 GalNAc had an R^2 value of 0.95. These data confirm that the di-HexNAc terminal epitope is terminating in a terminal 4-linked GalNAc. However, due to the loss of the anomeric configuration in the $^{0,2}A-H_2O$ fragment ion (Doohan et al. 2011), the laciNAc β -configuration needed to be confirmed using exoglycosidase digestion (see below). We could also verify that both HGM-1 and HGM-2 contained terminal GalNAc residues, using the lectin from *Wisteria floribunda* (Table I), previously used to detect laciNAc epitopes in HGMs. The amount of lectin binding to the HGM-2 was higher

compared with HGM-1 as indicated also in the LC-MS chromatogram (Figure 2) of the laciNAc type oligosaccharides.

In order to confirm that the linkage configuration of the terminal GalNAc, we used exo-*N*-acetylhexosaminidase digestion. With limitations on the amount of HGM available for analysis, we resorted using porcine gastric mucins (PGMs) as an alternative as we could identify the identical structure with an [M-H]⁻ ion of m/z 936 and RT 24.1 min (Supplementary data, Figures S1 and S2). Extensive investigation of exoglycosidase digestions showed that the structure was unaffected by β -hexosaminidases that remove terminal β 2,4- and 6-linked GlcNAc (*Saccharomyces cerevisiae*) and α -*N*-acetylgalactosaminidase (Figure 4). Treatment with β -*N*-acetylhexosaminidase from jack bean (JBH), which removes both β -linked GlcNAc and GalNAc residues, showed that a structure with a terminal β 1-6-linked GlcNAc was significantly digested within hours after initiating the digestion, whereas the [M-H]⁻ ion of m/z 936 isomer containing the 4-linked GalNAc required extended digestion, consistent with a β 1-4 configuration (Figure 4). The preferred removal of 6-linked over to 4-linked *N*-acetylhexosamine has been reported for this enzyme (Peracaula et al. 2003) and confirms, together with the evidence above, that the m/z 936 isomer contains a laciNAc moiety. Increased amount of the product Fuc α 1-2Gal β 1-3GalNAcol was detected in the LC-MS after

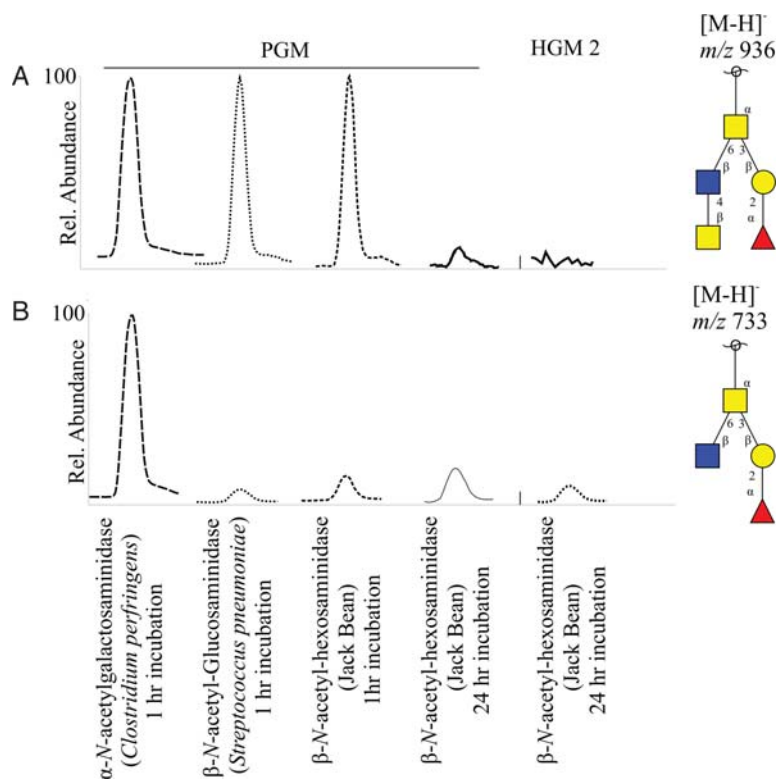


Fig. 4. Confirmation of the β -linkage configuration of the laciNAc moiety in the gastric Fuc α 1-2Gal β 1-3(GalNAc β 1-4GlcNAc β 1-6)GalNAcol structure by *N*-acetylhexosaminidases and LC-MS. Effect on the LC-MS intensity of the laciNAc containing structure Fuc α 1-2Gal β 1-3(GalNAc β 1-4GlcNAc β 1-6)GalNAcol isolated from PGM and HGM-2 (A) and the control Fuc α 1-2Gal β 1-3(GlcNAc β 1-6)GalNAcol from PGM and HGM-2 (B) as measured by LC-MS using various *N*-acetylhexosaminidases. Intensities are normalized against internal standard Fuc α 1-2Gal β 1-3(Fuc α 1-2Gal β 1-4GlcNAc β 1-6)GalNAcol. Key for symbols, see Figure 2.

the treatment having removed both *N*-acetylhexosamines of the C-6 branch on the GalNAcol (data not shown). The oligosaccharides from HGM-2 was also subjected to extended JBH treatment, and the lacdiNAc containing structure from HGM-2 with the $[M-H]^-$ ion of m/z 936 was also digested completely (Figure 4), as was its unglycosylated structure $[M-H]^-$ ion of m/z 790. These data confirm that the structure described in Figure 3 is indeed Fuc α 1-2Gal β 1-3(GalNAc β 1-4GlcNAc β 1-6)GalNAcol and is present in both human and pig gastric mucins.

Discussion

It has been shown that a GalNAc transferase (β 1,4-*N*-acetylgalactosaminyltransferase III) is present in human stomach capable of making the lacdiNAc epitope, with a specificity toward the core 2 type structures of *O*-linked oligosaccharides. The data from staining with the lectin from *W. floribunda*, binding to terminal GalNAc, indicated in the previous report that lacdiNAc is found on a secreted mucin present on the gastric mucosal lining (Ikehara et al. 2006). We believe that the identified lacdiNAc structures on MUC5AC detected in this report are the biological products from this enzyme. LacdiNAc as a terminal structure is a rare modification in mammals and is most frequently described in *N*-linked oligosaccharides (Green et al. 1985), but more recently it has also been identified on *O*-linked oligosaccharides from *zona pellucida* 3 (Dell et al. 2003). The only structural difference between lacdiNAc and lacNAc (Gal β 1-4GlcNAc β 1-) is the exchange of a C-2-OH to a C-2-NHAc on the outermost Gal residue. Although this looks like a minor difference, the consequences in the biosynthesis are dramatic. Whereas the lacNAc motif is used for the elongation of the oligosaccharides as well as decoration with fucose, sialic acid and sulfate residues, the lacdiNAc epitope appears to be a terminal motif resulting in fewer complex structures with limited decoration, possibly only with sulfate or sialic acid (Green et al. 1985; Marino et al. 2011). This novel terminal epitope may provide additional functionality into the interplay between gastric mucosa and gastric bacteria, where other terminal gastric epitopes such as the Le^b structure and sialyl Le^{x/a} being the target for *H. pylori* binding via BabA (Ilver et al. 1998) or SabA (Mahdavi et al. 2002) and the GlcNAc α 1- epitope (Kawakubo et al. 2004) that has been shown to have antibacterial properties. Additional lectin adhesins have been indicated to be present in the *H. pylori* genome, in the phylogenetic tree of OMPs containing BabA, BabB, BabC and SabA (Alm et al. 2000). Additional function and specificity of members of this family is unknown. Interaction with an epitope such as lacdiNAc is complementary to the Le^b epitope, since it would not be secretor-dependent and is present in the normal gastric tissue, while binding to SabA is dependent on the induction of sialylation due to *H. pylori*-induced gastritis. The significance of the presence of the lacdiNAc epitope on the gastric mucosal layer will need further investigation. The data in this report indicate that it may be involved in suppressing the binding to and the proliferation of *H. pylori*. This could either be via an active interaction of the epitope and the bacteria or an indirect effect, where the termination of lacdiNAc

influence the expression of Le^b and complex sialylated structures that are known to influence *H. pylori* bacterium's ability to establish and thrive in the gastric area.

Materials and methods

All materials were obtained from Sigma Aldrich (St Louis, MO) unless otherwise stated. The 18-m Ω water was produced using the MilliQ water purification system (Millipore, Billerica, MA).

Biological sample preparation

Oligosaccharides from PGM for hexosaminidase digestion and LC-MS were obtained from Sigma Aldrich or prepared from the antrum of the porcine stomach as per the method described earlier (Nordman et al. 1998). Human gastric specimens were obtained after informed consent and the approval of local ethics committee (Lund University Hospital, Lund, Sweden). The mucin sample HGM-1 was isolated from the antrum of the normal (normal as determined by a clinical pathologist) part of the stomach of a patient with gastric cancer, and HGM-2 was isolated from the junction between antrum and corpus from a patient undergoing obesity surgery. Frozen specimens were thawed in the presence of di-isopropyl phosphorofluoridate. Briefly, mucins were extracted in 6 M guanidinium chloride/5 mM Na₂EDTA/5 mM *N*-ethylmaleimide/10 mM sodium phosphate buffer, pH 6.5, and purified by density gradient centrifugation in CsCl/4 M guanidinium chloride as per the method described previously (Lindén et al. 2004). Samples were confirmed to contain MUC5AC with only small contribution of MUC6 (Skoog et al. 2011).

Binding and proliferation of *H. pylori* J99 to HGM

Binding of the *H. pylori* strain J99 to mucins was investigated by a microtiter-based assay. The 96-well polysorb plates (NUNC A/S, Roskilde, Denmark) were coated with the isolated HGM diluted in 4 M guanidine hydrochloride. The microtiter plates were washed three times in washing buffer between all ensuing steps. Unbound sites were blocked with a blocking reagent (Roche, Basel, Switzerland) containing 0.05% Tween. Biotinylated *H. pylori* J99 was added to the mucins and incubated in a bacterial shaker at 37°C for 2 h. The plates were washed three times and incubated for 1 h at room temperature with horseradish peroxidase-conjugated streptavidin. After further washings, the tetramethylbenzidine substrate was added. The reaction was stopped after 25 min with H₂SO₄ and the plates were read using a microplate reader at 450 nm.

The proliferation of *H. pylori* J99 when cultured on HGM-1 and HGM-2 was tested over a 60-h period of time. J99 was cultured in 96-well plates with either HGM-1 or HGM-2 in brain heart infusion broth supplemented with 10% fetal bovine serum. OD₅₆₀ measurements were recorded at regular intervals over a 60-h period. Any significant differences in the proliferation level compared with the J99 proliferation in the absence of mucins after the 60-h incubation period were recorded. Significantly lower proliferation compared with the control was considered as a negative result.

Release of *O*-linked oligosaccharides for LC-MS and *N*-acetylhexosaminidase digestion

The glyco-variant of MUC5AC isolated as HGM-1, HGM-2 and PGMs was blotted onto Immobillin P PVDF membranes (Millipore) and stained with DB71 and destained in 10% acetic acid in 40% ethanol. The *O*-linked oligosaccharides were released in 0.5 M NaBH₄ in 50 mM NaOH (Karlsson et al. 2004). The reaction was suspended with acetic acid and the samples were desalted with 60 µL of AG50WX8 cation exchange beads (Bio-Rad, Hercules, CA) packed in C18 zip tips (Millipore). Borate complexes were removed by repeated addition/evaporation with 1% acetic acid in methanol (100 µL for each addition). The released oligosaccharides were dissolved in water for introduction to the LC-MS.

LC-MS and LC-MSⁿ by CID of released oligosaccharides from HGM and PGM

The isolated MUC5AC oligosaccharides were analyzed by LC-MS and LC-MSⁿ using a 10 cm × 250 µm I.D. column containing 5-µm porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA) prepared in-house (Kenny et al. 2011). Oligosaccharides were eluted using a linear gradient from 0–40% acetonitrile over 40 min at a flow rate of 10 µL/min. The eluted oligosaccharides were detected in an ESI-IT MS (LTQ, Thermo Electron Corp., San Jose, CA) in the negative ion mode with a spray voltage of 3.5 kV. Air was used as a sheath gas and mass ranges were defined as per the specific structure to be analyzed. Specified ions were isolated for MSⁿ fragmentation by CID with the collision energy set to 30%. The data were manually interpreted following the guidelines as described (Karlsson et al. 2004) for the interpretation of *O*-linked oligosaccharides. Structural assignment was performed by the comparison of MS² spectra from isolated chromatographic peaks to structures identified in the UniCarb-DB glycomic database (Hayes et al. 2011). Identified structures in Table II and Supplementary data, Table SI have been submitted to UniCarb-DB (www.unicarb.org).

A GlcNAcβ1-4GlcNAc β1-4GlcNAc standard (Sigma Aldrich) and a GalNAcβ1-4-Gal standard (DextraUK, Reading, UK) were used to obtain the fragmentation spectra of a terminal β1-4-linked GalNAc and a β1-4-linked GlcNAc. The standards were prepared in water to a concentration of 10 µg/µL and analyzed by LC-MS with an in-source fragmentation of 30%. The ^{0,2}A₂ fragment from chitotriose and ^{0,2}A₂ fragment from GalNAc-β1-4-Gal were isolated with an isolation width of ±3 mass units for MS³ fragmentation with collision energy of 30%. Specific structures from HGM-2, identified as having an unknown terminal HexNAc structure by the presence of the diagnostic fragment ion *m/z* 610, were also isolated for MS³ fragmentation.

Hexosaminidase treatment of PGM and HGM-2

The LC-MS of oligosaccharides from PGM (Nordman et al. 1998) after reductive β-elimination showed a selective reaction monitoring chromatogram of three parent ions that fragmented into the specific Z-fragment of *m/z* 610 (Supplementary data, Figure S1). This fragment ion is indicative for a HexNAc-HexNAc-HexNAc species and the nature is described in Figure 3A. These parent ions also showed the diagnostic ^{0,2}A

fragment of *m/z* 304 from 4-linked HexNAc (Supplementary data, Figure S2 for the [M–H][–] ion of *m/z* 936). The [M–H][–] ion of the *m/z* 936 structure was shown by RT and LC-MS² to be identical to the human lacdiNAc containing species in HGM-2 and was used for exoglycosidase digestion and LC-MS to confirm the linkage of the terminal HexNAc.

O-Linked oligosaccharides (10 µg) released from PGM and HGM-2 were treated with 0.1 U of a JBH (Prozyme, Hayward, CA), β-*N*-acetylglucosaminidase (Prozyme from *S. pneumonia*) or α-*N*-acetylgalactosaminidase from *Clostridium perfringens* (R&D Systems, Minneapolis, MN) in 10 µL of the enzymes buffer. The reaction was incubated at 37°C for 1 h (or 24 h for JBH) and stopped by the addition of 1.0 M HCl. Digested and non-digested oligosaccharides were cleaned with 0.6 mL of HyperCarb® PGC (ThermoFisher Scientific, Waltham, MA) packed in C18 zip tips and analyzed by LC-MS and LC-MS². The relative intensities of oligosaccharides corresponding to a GlcNAcβ1-6(Fucα1-2Galβ1-3)GalNAc-ol ([M–H][–] ion of *m/z* 733) and GlcNAcβ1-GlcNAcβ1-6(Fucα1-2Galβ1-3)GalNAc-ol (*m/z* 936) were compared with the relative intensity of an internal standard of Fucα1-2Galβ1-4GlcNAcβ1-6(Fucα1-2Galβ1-3)GalNAc-ol (*m/z* 1041). The intensity of this particular oligosaccharide remains unchanged after enzymatic digestion).

Lectin and antibody probing of HGM-1, HGM-2 and PGM

HGM and PGM were transferred to PVDF membrane with the amount of mucin corresponding to 5 µg of carbohydrate content. Bovine fetuin was also transferred and used as a negative control. The membranes were blocked with 1% bovine albumin in Tris-buffered saline with 0.05% Tween-20 (TBST). Blots were probed with a biotinylated lectin from *W. floribunda* (Vector Laboratories, Burlingame, CA) diluted to 1:10,000 in TBST or mouse anti-Le^b mono-clonal antibodies (2-25LE, Abcam, Cambridge) diluted to 1:1000 in TBST for 3 h incubation at room temperature. After probing, the blot was washed five times with TBST for 5 min. The blots were incubated in horseradish peroxidase conjugated streptavidin diluted to 1:8000 in blocking solution for blots probed with lectin from *W. floribunda* and with horseradish peroxidase conjugated to goat anti-mouse IgG (P0161, Dako, Glostrup, Denmark) diluted to 1:20,000 in blocking solution for blots probed with the Le^b antibody for 60 min at room temperature. The blots were further washed for five times with TBST and once with TBS. The blot was developed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific). Signal intensity was measured using ImageJ software (National Institute of Health, Bethesda, MA).

Supplementary data

Supplementary data for this article is available online at <http://glycob.oxfordjournals.org/>.

Acknowledgements

The antral mucin from porcine stomach was a gift from Prof. Ingemar Carlstedt and Dr Henrik Nordman, Lunds University.

Dr Catherine A Hayes is acknowledged for her help in uploading the MS data to UniCarb-DB.

Funding

This work was supported by the Swedish Research Council (621-2010-5322 and K20008-58X-20693-01-4), The Swedish Foundation for International Cooperation in Research and Higher Education, Åke Vibergs Foundation and the Swedish Cancer Foundation. The mass spectrometer was obtained by a grant from the Swedish Research Council (342-2004-4434).

Conflict of interest

None declared.

Abbreviation

BabA, blood group antigen binding adhesin; HMG, human gastric mucin; JBH, jack bean β -*N*-acetylhexosaminidase; lactiNac, di-*N*-acetylglucosamine; OMP, outer membrane protein; PGC, porous graphitized carbon; PGC, porous graphitized carbon; PGM, porcine gastric mucin; RT, retention time; SabA, sialic acid-binding adhesin; SIC, selected ion chromatogram; TBST, Tris-buffered saline with Tween-20; CID, collision induced dissociation; Fuc, Fucose; GalNac, *N*-acetylgalactosamine; GlcNac, *N*-acetylglucosamine; Hex, hexose; HexNac, *N*-acetylhexosamine; LC-MS, liquid chromatography-mass spectrometry; PVDF, polyvinylidene fluoride

References

- Alm RA, Bina J, Andrews BM, Doig P, Hancock REW, Trust TJ. 2000. Comparative genomics of *Helicobacter pylori*: Analysis of the outer membrane protein families. *Infect Immun.* 68(7):4155–4168.
- Cooke CL, An HJ, Kim J, Canfield DR, Torres J, Lebrilla CB, Solnick JV. 2009. Modification of gastric mucin oligosaccharide expression in rhesus macaques after infection with *Helicobacter pylori*. *Gastroenterol.* 137(3):1061–1071.
- Dell A, Chalabi S, Easton RL, Haslam SM, Sutton-Smith M, Patankar MS, Lattanzio F, Panico M, Morris HR, Clark GF. 2003. Murine and human zona pellucida 3 derived from mouse eggs express identical O-glycans. *Proc Natl Acad Sci USA.* 100(26):15631–15636.
- Doohan RA, Hayes CA, Harhen B, Karlsson NG. 2011. Negative ion CID fragmentation of O-linked oligosaccharide aldoses—charge induced and charge remote fragmentation. *J Am Soc Mass Spectrom.* 22(6):1052–1062.
- Green ED, van Halbeek H, Boime I, Baenziger JU. 1985. Structural elucidation of the disulfated oligosaccharide from bovine lutropin. *J Biol Chem.* 260(29):15623–15630.
- Hayes CA, Karlsson NG, Struwe WB, Lisacek F, Rudd PM, Packer NH, Campbell MP. 2011. UniCarb-DB: A database resource for glycomic discovery. *Bioinformatics.* 27(9):1343–1344.
- Ikehara Y, Sato T, Niwa T, Nakamura S, Gotoh M, Ikehara SK, Kiyohara K, Aoki C, Iwai T, Nakanishi H, et al. 2006. Apical Golgi localization of *N,N*-diacetyllactosylamine synthase, β 4GalNac-T3, is responsible for LactiNac expression on gastric mucosa. *Glycobiology.* 16(9):777–785.
- Ilver D, Arnqvist A, Ögren J, Frick IM, Kersulyte D, Incecik ET, Berg DE, Covacci A, Engstrand L, Borén T. 1998. *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science.* 279(5349):373–377.
- Karlsson NG, Schulz BL, Packer NH. 2004. Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J Am Soc Mass Spectrom.* 15(5):659–672.
- Kawakubo M, Ito Y, Okimura Y, Kobayashi M, Sakura K, Kasama S, Fukuda MN, Fukuda M, Katsuyama T, Nakayama J. 2004. Natural antibiotic function of a human gastric mucin against *Helicobacter pylori* infection. *Science.* 305(5686):1003–1006.
- Kenny DT, Issa SMA, Karlsson NG. 2011. Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation. *Rapid Commun Mass Spectrom.* 25(18):2611–2618.
- Kobayashi M, Lee H, Nakayama J, Fukuda M. 2009. Roles of gastric mucin-type O-glycans in the pathogenesis of *Helicobacter pylori* infection. *Glycobiology.* 19(5):453–461.
- Lindén S, Borén T, Dubois A, Carlstedt I. 2004. Rhesus monkey gastric mucins: Oligomeric structure, glycoforms and *Helicobacter pylori* binding. *Biochem J.* 379:765–775.
- Lindén S, Nordman H, Hedenbro J, Hurtig M, Borén T, Carlstedt I. 2002. Strain- and blood group-dependent binding of *Helicobacter pylori* to human gastric MUC5AC glycoforms. *Gastroenterology.* 123(6):1923–1930.
- Mahdavi J, Sondén B, Hurtig M, Olfat FO, Forsberg L, Roche N, Ångström J, Larsson T, Teneberg S, Karlsson K-A, et al. 2002. *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *Science.* 297(5581):573–578.
- Marino K, Lane JA, Abrahams JL, Struwe WB, Harvey DJ, Marotta M, Hickey RM, Rudd PM. 2011. Method for milk oligosaccharide profiling by 2-aminobenzamide labeling and hydrophilic interaction chromatography. *Glycobiology.* 21(10):1317–1330.
- Marshall BJ, Warren JR. 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic-ulceration. *Lancet.* 1(8390):1311–1315.
- McGuckin MA, Lindén SK, Sutton P, Florin TH. 2011. Mucin dynamics and enteric pathogens. *Nat Rev Microbiol.* 9(4):265–278.
- Menaker RJ, Sharaf AA, Jones NL. 2004. *Helicobacter pylori* infection and gastric cancer: Host, bug, environment, or all three? *Curr Gastroenterology Rep.* 6(6):429–435.
- Newman W, Kabat EA. 1976. Immunochemical studies on blood groups. Structures and immunochemical properties of nine oligosaccharides from B-active and non-B-active blood group substances of horse gastric mucosae. *Arch Biochem Biophys.* 172(2):535–550.
- Nordman H, Davies JR, Carlstedt I. 1998. Mucus glycoproteins from pig gastric mucosa: Different mucins are produced by the surface epithelium and the glands. *Biochem J.* 331(3):687–694.
- Nordman H, Davies JR, Lindell G, de Bolos C, Real F, Carlstedt I. 2002. Gastric MUC5AC and MUC6 are large oligomeric mucins that differ in size, glycosylation and tissue distribution. *Biochem J.* 364:191–200.
- Peracaula R, Tabarés G, Royle L, Harvey DJ, Dwek RA, Rudd PM, de Llorens R. 2003. Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins. *Glycobiology.* 13(6):457–470.
- Sakamoto J, Watanabe T, Tokumaru T, Takagi H, Nakazato H, Lloyd KO. 1989. Expression of Lewis-A, Lewis-B, Lewis-X, Lewis-Y, sialyl-Lewis-A, and sialyl-Lewis-X blood-group antigens in human gastric-carcinoma and in normal gastric tissue. *Cancer Res.* 49(3):745–752.
- Skoog EC, Lindberg M, Linden SK. 2011. Strain-dependent proliferation in response to human gastric mucin and adhesion properties of *Helicobacter pylori* are not affected by co-isolated *Lactobacillus* sp. *Helicobacter.* 16(1):9–19.
- Styer CM, Hansen LM, Cooke CL, Gundersen AM, Choi SS, Berg DE, Benghezal M, Marshall BJ, Peek RM, Boren T, et al. 2010. Expression of the BabA adhesin during experimental infection with *Helicobacter pylori*. *Infect Immun.* 78(4):1593–1600.

Supplementary Table I Structures identified on the HMG-1 samples both binding to and inducing proliferation of the *H. pylori* strain Y99

[M-H]⁻	Structures^a
384 ⁻	GalNAc-ol Galβ1-3
513 ⁻	NeuAcα2-6 GalNAcol
530 ⁻	GalNAc-ol Fuca1-2Galβ1-3
587 ⁻	GlcNAcβ1-6 GalNAc-ol Galβ1-3
587 ⁻	GalNAc-ol GlcNAcα1-4Galβ1-3
587 ⁻	GalNAc-ol Galβ1-4GlcNAcβ1-3
675 ⁻	NeuAcα2-6 GalNAcol Galβ1-3
675 ⁻	GalNAcol NeuAcα2-Galβ1-3
733 ⁻	GalNAc-ol Galβ1-3(Fuca1-3/4)GlcNAcβ1-3
733 ⁻	GalNAc-ol Fuca1-2Galβ1-3GlcNAcβ1-3
733 ⁻	GalNAc-ol GalNAcα1-3(Fuca1-2)Galβ1-3
733 ⁻	GalNAc-ol Fuca1-2Galβ1-4GlcNAcβ1-3
733 ⁻	GlcNAcβ1-6 GalNAc-ol
733 ⁻	GalNAc-ol Fuca1-2Gal β1-3 Galβ1-4GlcNAcβ1-6
749 ⁻	GalNAc-ol ^b Galβ1-3
749 ⁻	Galβ1-4GlcNAcβ1-3Galβ1-3 GalNAc-ol
790 ⁻	<u>GalNAcβ1-4GlcNAcβ1-6</u> GalNAc-ol Galβ1-3
790 ⁻	GlcNAcβ1-6 GalNAc-ol
790 ⁻	GalNAc-ol GlcNAcα1-4Galβ1-3
878 ⁻	GlcNAcβ1-6

		GalNAc-ol
	NeuAc2-Galβ1-3	
895 ⁻	Galβ1-4/3(Fuca1-3/4)GlcNAcβ1-6	GalNAc-ol
		Galβ1-3
895 ⁻		GalNAc-ol
	Galα1-3(Fuca1-2)Galβ1-3GlcNAcβ1-3	
895 ⁻	Fuca1-2Galβ1-4GlcNAcβ1-6	GalNAc-ol
		Galβ1-3
895 ⁻	Galβ1-3GlcNAcβ1-6	GalNAc-ol
	Fuca1-2Galβ1-3	
936 ^{-b}		GalNAc-ol
	GalNAcα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-3	
936 ^{-b}		GalNAc-ol
	Fuca1-2GlcNAcβ1-3Galβ1-3	
936 ^{-b}	Fuca1-2Galβ1-4GlcNAcβ1-6	GalNAc-ol
		GlcNAcβ1-3
936 ^{-b}		GalNAc-ol
	<u>GalNAcβ1-4GlcNAcβ1-6</u>	
936 ^{-b}		GalNAc-ol
	Fuca1-2Gal β1-3	
952 ⁻	GlcNAcα1-4Galβ1-4GlcNAcβ1-Galβ1-3	GalNAc-ol
966 ⁻	NeuAcα2-6	GalNAcol
	NeuAcα2-Galβ1-3	
1040 ⁻	NeuAcα2- { Galβ1-GlcNAcβ1-6	GalNAc-ol
		Galβ1-3
1040 ⁻	NeuAcα2- { Galβ1-GlcNAcβ1-6	GalNAc-ol
		Galβ1-3
1041 ⁻	Fuca1-2Galβ1-4GlcNAcβ1-6	GalNAc-ol
	Fuca1-2Galβ1-3	
1057 ⁻	Galα1-3(Fuca1-2)Galβ1-3GlcNAcβ1-6	GalNAc-ol
		Galβ1-3
1057 ⁻	Galβ1-GlcNAcβ1-6	GalNAc-ol
	Galα1-3(Fuca1-2)Galβ1-3	

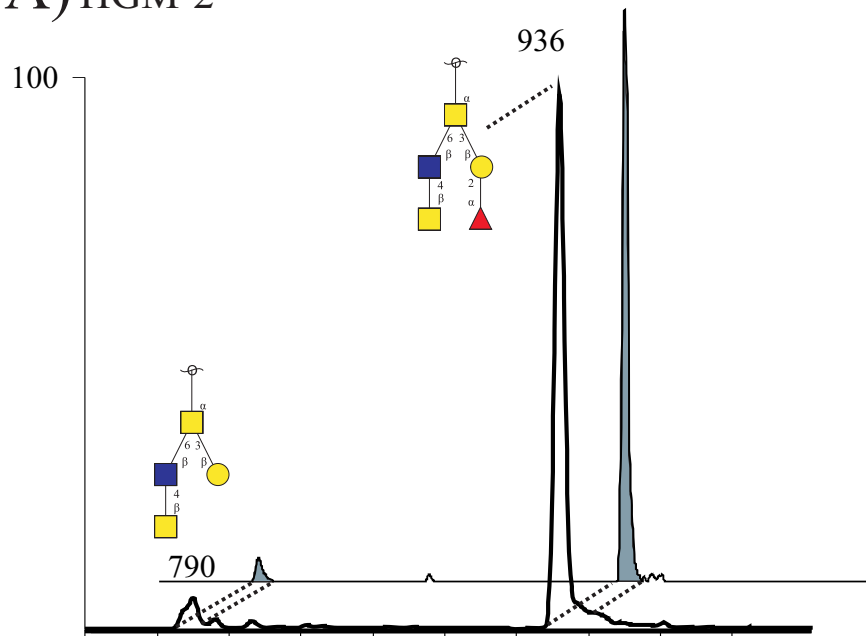
1098 ⁻		Galβ1-GlcNAcβ1-6 GalNAc-ol
		GalNAcα1-3(Fuca1-2)Galβ1-3
1098 ⁻		Fuca1-2Galβ1-4GlcNAcβ1-6 GalNAc-ol
		GlcNAc1-Galβ1-3
1098 ⁻		GalNAcα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-6 GalNAc-ol
		Galβ1-3
1186 ⁻	NeuAcα2-	Galβ1-4/3(Fuca1-3/4)GlcNAcβ1-6 GalNAc-ol
		Gal β1-3
1186 ⁻		Fuca1-2Gal β1-4GlcNAc β1-6 GalNAc-ol
		NeuAcα2-Gal β1-3
1186 ⁻		NeuAcα2-Galβ1-GlcNAcβ1-6 GalNAc-ol
		Fuca1-2Galβ1-3
1203 ⁻		Galα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-6 GalNAc-ol ^a
		Fuca1-2Galβ1-3
1244 ⁻		GalNAcα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-6 GalNAc-ol
		Fuca1-2Galβ1-3
665 ²⁻	2×NeuAcα2-	Galβ1-GlcNAcβ1-6 GalNAc-ol
		Galβ1-3
1332 ⁻		NeuAcα2-Galβ1-4/3(Fuca1-3/4)GlcNAcβ1-6 GalNAc-ol
		Fuca1-2Galβ1-3
1365 ⁻		Galα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-6 GalNAc-ol
		Galβ1-3(Fuca1-2)Galβ1-3
1406 ⁻		Galα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-6 GalNAc-ol
		GalNAcα1-3(Fuca1-2)Galβ1-3
1447 ⁻		GalNAcα1-3(Fuca1-2)Galβ1-4GlcNAcβ1-6 GalNAc-ol
		GalNAcα1-3(Fuca1-2)Galβ1-3
1477 ⁻		NeuAcα2-Galβ1-4/3(Fuca1-3/4)GlcNAcβ1-6 GalNAc-ol
		NeuAcα2-Galβ1-3

^aSequences identified by comparison of MSⁿ spectra from UniCarb-DB(www.unicarb-

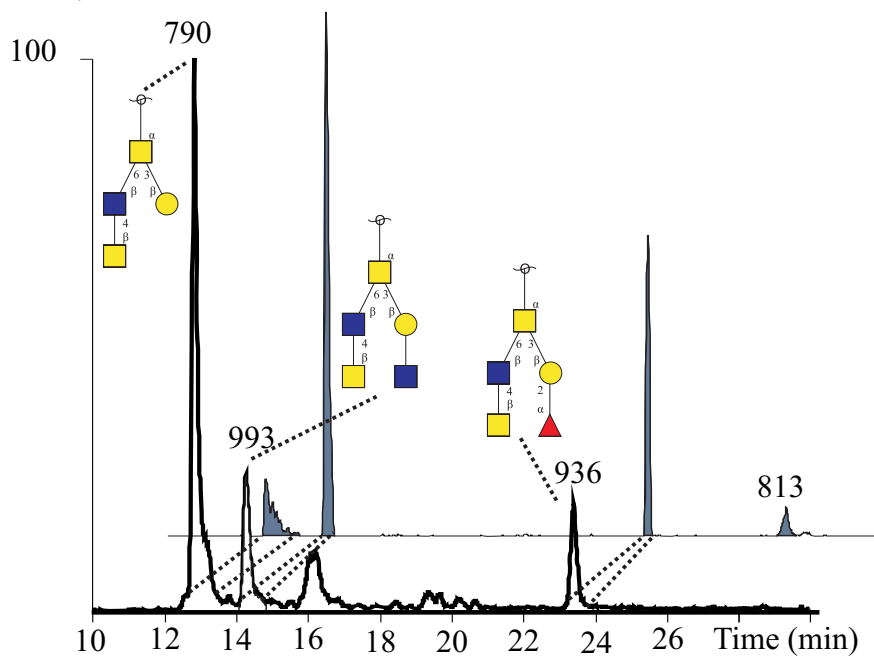
DB.org). Linkage configuration and positions are based on core 1, 2 and 3 series of extension. Type 2 configuration Gal β 1-4GlcNAc β 1- is based on detection of C-4 linkage specific $^{0,2}A$ fragment ions and type 1 is based on the absence of these ions. Terminal epitope GlcNAc β 1-4GlcNAc is assumed based on evidence in this report and presence of the $^{0,2}A$ fragment ion of m/z 304 in HexNAc-HexNAc sequences confirming linkage position. Similarly the GlcNAc α 1-4Gal epitope is assumed and based on the presence of the $^{0,2}A$ fragment ion of m/z 304 of HexNAc-Hex sequences. Fuc configuration is assumed to be α 1-2 if present in a blood group H sequence and α 1-3/4 if linked to GlcNAc. GalNAc α 1-3 and Gal α 1-3 is assumed in blood group AB sequences. Hex always assumed to be Gal and HexNAc to be GlcNAc except in blood group A. Reducing end HexNAc-ol assumed to be GalNAc-ol. Bolded and underlined sequences correspond to C-6 terminating lacdiNAc.

^bRefer to number 1-4 in order of elution in Figure 2A for [M-H]⁻ ion of m/z 936 of HGM1.

A) HGM-2



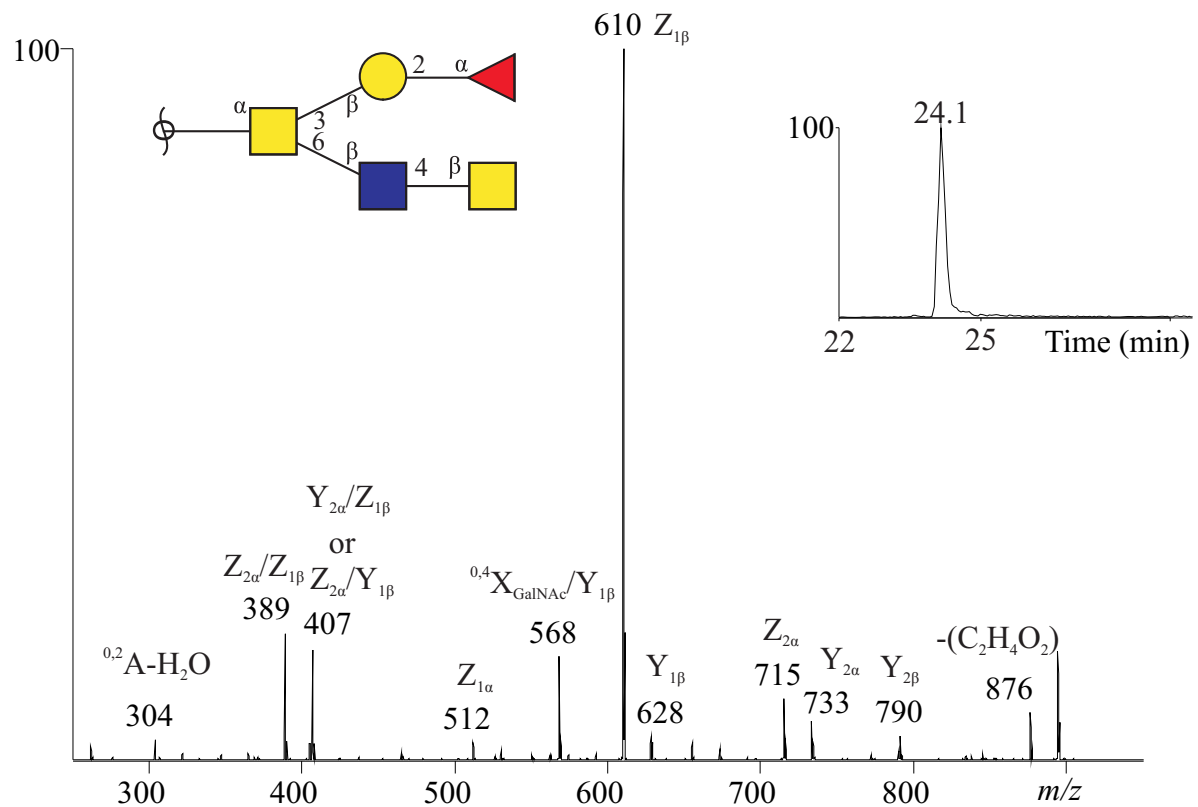
B) PGM



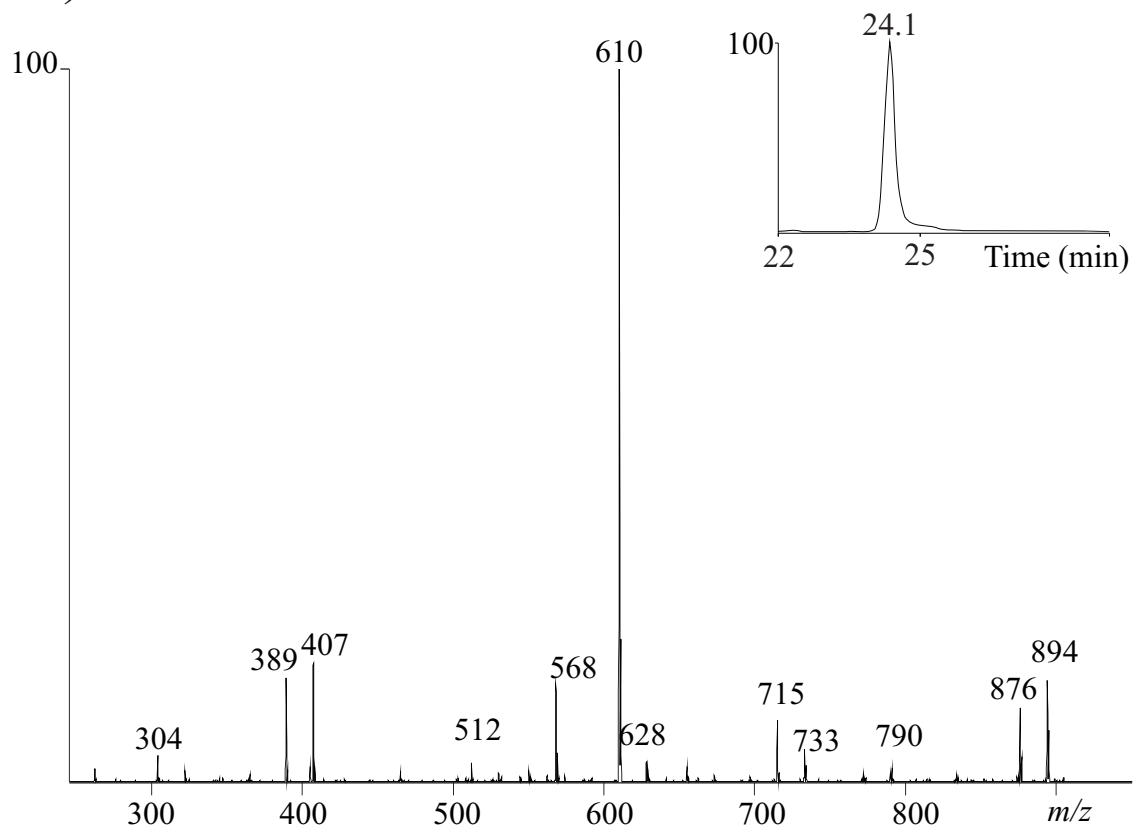
Supplementary Figure 1. Identification of lactiNAc in human and porcine gastric mucin

Selected ion MS² chromatograms of the pronounced Z fragment GalNAc β 1-4GlcNAc β 1-6GalNAcol of *m/z* 610 (grey) combined with combined SIMs of *m/z* 790, 936 and 993 of HGM-2 (A) and antral glandular porcine mucin (B), showing the presence of two and three lactiNAc containing components in these samples, respectively. The component *m/z* 790 and *m/z* 936 has identical retention time in the human and the porcine gastric samples. Key for symbols see Figure 2.

A) MS² of [M-H]⁻ *m/z* 936 from HGM2



B) MS² of [M-H]⁻ *m/z* 936 from PGM



Supplementary Figure 2. Comparison of LC-MS² spectrum of the LacdiNAc containing structure from Figure 2A from human gastric mucin (HGM1) A) and B) porcine gastric mucin (PGM).

MS² spectra for the lacdiNAc containing structure $Fuc\alpha 1-2Gal\beta 1-3(GalNAc\beta 1-4GlcNAc\beta 1-6)GalNAcol$ with an $[M-H]^{-}$ -ion of *m/z* 936 spectra for HGM1 and PGM showing that they have the same MS². The inserts shows a selected ion chromatogram with the same retention time in each sample, confirming that the same structure is present in both HGM and PGM.