Optimisation of glycomics analysis and the effect of inflammation on protein glycosylation in the central nervous system

Christopher David Ashwood

BSci Forensic Science, Bond University
MRes Molecular Sciences, Macquarie University

ARC Centre of Nanoscale Biophotonics

Glyco@MQ

Department of Molecular Sciences

Macquarie University, Sydney, Australia

A thesis presented for the degree of

Doctor of Philosophy

On

5th August 2018

STATEMENT OF ORIGINALITY

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. This work has not been submitted for a higher degree to any other university or institution.

The research presented in this thesis was given Exempt Dealing status by the Macquarie University Biosafety Committee: "Interacting ligands of the TLR4-mediated signalling pathway and the molecular consequences of its activation" [5201500760].

Some of the research presented in this thesis (Chapter 5) was approved by the University of Adelaide Animal Ethics committee: “Pain caused by the classical and non-classical brain cells” [M-2013-227] and was noted by the Macquarie University Animal Ethics committee (0646 - 520180646866 - Ashwood - Pain caused by the classical and non-classical brain cell)

Christopher Ashwood

TABLE OF CONTENTS

LIST OF PUBLICATIONS, PRESENTATIONS AND AWARDS

Publications

This thesis is based on the following original publications, which are referred to in the text by Roman numerals. Original publications are reproduced with permission from their copyright holders.

Chapter 1

I. Everest-Dass, A. V, Moh, E.S.X.*, **Ashwood, C.*,** Shathili, A.M.M.*, Packer, N.H.: Human disease glycomics: technology advances enabling protein glycosylation analysis – part 1. Expert Rev. Proteomics. 15, 165-182 (2017).

II. Everest-Dass, A. V., Moh, E.S.X.*, **Ashwood, C.*,** Shathili, A.M.M.*, Packer, N.H.: Human disease glycomics: technology advances enabling protein glycosylation analysis–part 2. Expert Rev. Proteomics. 15, 341-352 (2018)

Chapter 3

III. **Ashwood, C.,** Abrahams, J.L., Nevalainen, H., Packer, N.H.: Enhancing structural characterisation of glucuronidated O -linked glycans using negative mode ion trap-HCD mass spectrometry. Rapid Commun. Mass Spectrom. 30, 851–858 (2017)

Chapter 5

IV. **Ashwood, C.,** Lin, C.-H., Thaysen-Andersen, M., Packer, N.H.: Discrimination of Isomers of Released *N*- and *O*-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. J. Am. Soc. Mass Spectrom. doi: 10.1007/s13361-018-1932-z. [Epub ahead of print] (2018)

* Equal contributors

Presentations

- **Skyline User Group Meeting Oral Presentation** (San Diego, USA) Applications of Skyline for automated profiling of cellular protein glycosylation (**C Ashwood**)
- **Human Proteome Organisation Oral Presentation (Multiplex technologies session)** (Dublin, Ireland 2017) Open glycomics: an open-access platform for automated glycan identification and quantitation. (**C Ashwood**, CH Lin, M Thaysen-Andersen and NH Packer)
- **American Society for Mass Spectrometry Poster Presentation** (Indianapolis, USA 2017) Skyline-based quantitative *N*-glycomics reveals altered expression of paucimannosidic *N*-glycans of mouse brain tissues upon neuropathic pain induction (**C**

Ashwood, V Staikopoulos, AV Everest-Dass, M Thaysen-Andersen, MR Hutchinson, NH Packer)

- **Australasian Protein Society Oral Presentation (Glycomics and glycoproteomics session)** (Lorne, Australia 2017) Revealing glycan structural epitopes using a simplified LC-MS/MS data acquisition and analysis platform for glycomics (**C Ashwood**, CH Lin, M Thaysen-Andersen and NH Packer)
- **Human Proteome Organisation ThermoFisher Industry Scientific Seminar Oral Presentation** (Taipei, Taiwan 2016) Unlocking the Low Mass Range with Trap-HCD for Glycan Analysis (**C Ashwood**, JL Abrahams, NH Packer)

Awards

- **Australian Glycosciences Symposium Poster Award** ($250 AUD, 2018) Taming the beast: Standardising porous graphitised carbon based LC-MS glycomics. (**C Ashwood**, NH Packer)
- **APS HUPO Student Travel Award** ($500 AUD, 2017) Open glycomics: an open-access platform for automated glycan identification and quantitation. (**C Ashwood**, CH Lin, M Thaysen-Andersen and NH Packer)
- **HUPO Poster Award** ($1000 USD, 2016) Improving confidence in glycan structural characterisation using alternative CID fragmentation (**C Ashwood**, JL Abrahams and NH Packer)
- **APS Poster Award** ($100 AUD, 2015) Novel *O*-glycans containing acidic sugars on fungal secreted protein and their characterisation using PGC-LC-ESI-MS/MS (**C. Ashwood**, H Nevalainen and NH Packer)
- **Macquarie Postgraduate Research Fund Award** ($4757 AUD, 2017) Attendance at ASMS 2017 and a three week lab visit to the Costello research group (Boston University).
- **Australian Postgraduate Award** (2015- 2018)

Supervisor

Prof. Nicolle H. Packer
Department of Molecular Sciences
Faculty of Science and Engineering
Macquarie University, Australia

Associate Supervisor

Prof. Mark Molloy
Australian Protein Analysis Facility
Department of Molecular Sciences
Faculty of Science and Engineering
Macquarie University, Australia

ACKNOWLEDGEMENTS

I would like to thank Macquarie University and the ARC Centre of Nanoscale Biophotonics for providing me with opportunities for professional development and being able to perform and present the research I have been a part of for the past 3 years. Without their support, I would not have learnt as much over the course of my PhD or met many expert glycomics researchers around the world. I would also like to thank the Australian Government for funding me to focus on my research through the Australian Postgraduate Award scholarship.

All members of the Glyco groups at Macquarie University, past and present, have in some way shaped who I am as a scientist and my research so I am very grateful to have been able to spend time with you all and learn about each of your research topics. I would also like to thank the members of the Hutchinson Group (Vicky, Sanam, Jacob and Mark) for introducing me to the field of chronic pain, providing mouse tissue samples and taking a chance with mass-spectrometry.

Thank you to Cathy, Deborah and Kevin for hosting me at Boston University, allowing me to learn about permethylated glycan analysis and various mass-spectrometric methods that I have not previously encountered. I would also like to thank members of the Molloy Group (Mark, Matt and Christoph) for their contributions towards my understanding of proteomics and development of targeted proteomic assays.

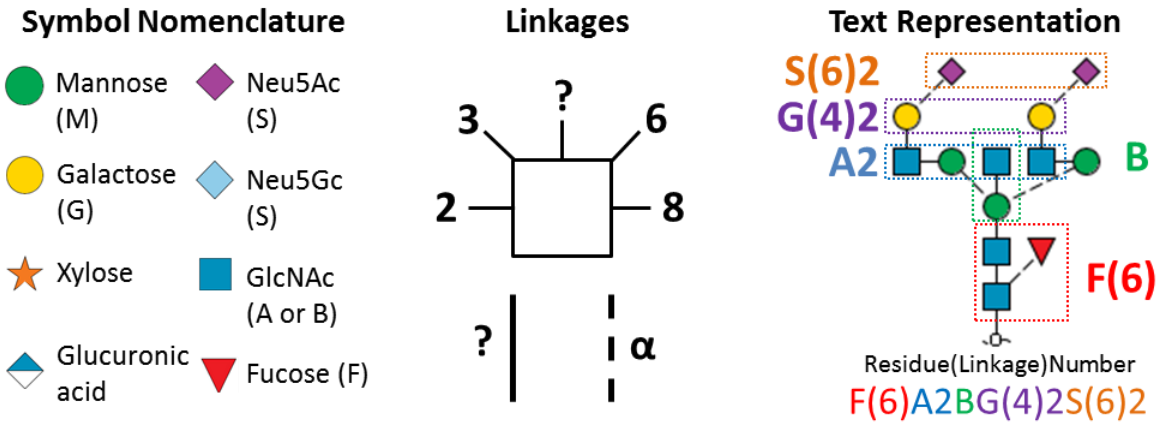
I would like to thank Jodie, Matthew and Paul for their immensely valuable training and advice. Your scientific knowledge and career advice will not be forgotten and I will strive to one-day mean the same to a mentee as you all have to me. I would also like to thank Arun, Edward, Abdulrahman, Chi-Hung and Morten for their direct contributions to a variety of projects from which have provided a great opportunity to learn more about you all and your valuable expertise.

Thank you to my supervisor Nicki Packer for her guidance, open-mindedness and support throughout my PhD, as my project changed from the initial plans into the contents of this thesis. I have had very good opportunities as a result of all those mentioned above but outside of the science, I would like to thank my family and Heather for providing their immense support which has allowed me to reach this milestone.

SYMBOLS AND NOMENCLATURE

Glycan nomenclature

To describe glycans by text, all *N*-glycans detected have, at minimum, two core GlcNAcs; F at the start of the abbreviation indicates a core (α1-6) fucose linked to the innermost GlcNAc; Mx, number (x) of mannose on core GlcNAcs; Ax, number of antenna (GlcNAc) on trimannosyl core; A2, biantennary with both GlcNAcs as β1-2 linked; A3, triantennary with a GlcNAc linked β1-2 to both mannose and a third GlcNAc linked β1-4 to the α1-3 linked mannose; A3', triantennary with a GlcNAc linked β1-2 to both mannose and a third GlcNAc linked β1-6 to the α1-6 linked mannose; A4, GlcNAcs linked as A3 with additional GlcNAc β1-6 linked to α1-6 mannose; B, bisecting GlcNAc linked β1-4 to α1-3 mannose; Gx, number (x) of β1-4 linked galactose on the antenna; Sx, number (x) of sialic acids linked to galactose.



To describe glycan structures visually, coloured shapes have been employed with the shape representing a monosaccharide exact mass (i.e. *N*-acetyl hexosamine is a square, hexose is a circle, deoxyhexose is a triangle) and the colour representing the monosaccharide stereochemistry (i.e. blue is glucose-type, yellow is galactose-type). The angles (e.g.. 0° is linked to the 2nd carbon of the attached monosaccharide, 45° is linked to the 3rd carbon of the attached monosaccharide) and lines (solid means known linkage anomericity of a β linkage, while dashed means α linkage) between these monosaccharides as per the Symbol Nomenclature For Glycans (SNFG)[1]. As drawing an ambiguous structure involves a solid line (typically indicates β linkage) in the 1,4 linkage position, we have erred on the side of caution by designating solid and/or vertical lines as unknown. We hope further advances in the SNFG can address this challenge.

Glycan structures shown in the thesis have all been created using GlycoWorkBench version 2.1 (available from https://code.google.com/archive/p/glycoworkbench/) as per SNFG with angled linkage placement (Oxford) in compact view and exported with a reducing end indicator. Linkages are drawn in the default 1-4 position are designated unknown linkage unless specified otherwise.

ABSTRACT

Glycomics aims to comprehensively study the entire complement of sugars found in an organism and ideally needs to detect and quantify each individual glycan structure present. Using liquid chromatography in tandem with mass-spectrometry allows separation, characterisation and quantitation of released glycan structures from biological samples, improving our understanding of the role glycosylation plays in life. For the nervous system, which is the most essential signalling system in the human body, the role of protein glycosylation is still largely not understood.

The application of liquid chromatography with mass-spectrometry for glycan analysis has been frequently performed over the past twenty years but the analysis of generated data is largely manual. Furthermore, there is no established workflow for the quantitation on glycans, independent of instrument manufacturer. Addressing this gap, we have established or optimised methods for obtaining glycan data from LC-MS systems, improved the characterisation of glycans using reproducible fragmentation methods and retention time normalisation and automated the assignment of glycan structure in challenging applications.

Toll-like receptor 4 is one of the research targets for understanding the signalling pathways involved in the human pain response due to its role as a membrane protein which mediates inflammation to both foreign and endogenous molecules. As glycosylation plays a role in cell-to-cell signalling, it is expected, but not yet established, that the activation of the toll-like receptor 4 pathway will result in changes to the cell surfaces involved in transmitting the signals involved in pain transduction. Furthermore, the abundance of toll-like receptor 4 in normal and activated signalling states has not yet been understood, despite the importance of receptor abundance in understanding a signalling pathway.

We have approached this signalling pathway with both *in vivo* and *in vitro* models observing up-regulation of core-fucosylation on proteins secreted by cell lines and identified a positive correlation of the paucimannose glycan class on specific brain regions to increased pain sensitivity. In addition, we have described the detection of toll-like receptor 4 in published studies and developed targeted methods to detect it and its partner, myeloid-differentiation factor 2.

CHAPTER 1 – INTRODUCTION

1.1 IDENTIFYING AND COUNTING BIOMOLECULES

If biochemistry is the study of chemical processes related to living organisms, analytical biochemistry is the separation, identification and quantification of the biomolecules related to these chemical processes. As a result of research efforts in analytical biochemistry, we can obtain characterisation and quantitation of the molecules involved in biological systems. This serves as a useful foundation upon which biological perturbations can be studied to understand the cause and effect of disease. Through increased understanding of disease, prevention, diagnosis and treatment can be more readily achieved.

The characterisation and quantitation of biomolecules relevant to disease has been made easier with time due to advances in the fields of physics and chemistry. The awarding of the Nobel prize in Chemistry to Fenn[2], Tanaka[3] and Wuthrich[4] in 2002 is one example of the recognition of these advancements, for the development of methods for identification and structure analyses of biological macromolecules. These advancements were also the result of earlier researchers such as Dehmelt and Paul (1989 Nobel Prize in Physics[5]) and Ernst (1991 Nobel Prize in Chemistry[6]) for development of the ion trap technique and developments of the methodology of high resolution nuclear magnetic resonance (NMR) spectroscopy.

Since the invention of these techniques and the discovery of these biomolecule classes, further investment in the life sciences has stimulated the growth of instrument speeds, accuracy and throughput and has correspondingly been used to answer an ever-increasing number of questions regarding biochemistry. One example of these instruments is mass-spectrometry, which detects and quantifies ions based on their mass-to-charge ratio, a feature that can be exploited for virtually all biomolecules[7].

1.2 DEVELOPMENT OF BIOMOLECULE CLASS-SPECIFIC IDENTIFICATION METHODS AND FIELDS

The diversity in chemical structure and size of biomolecules requires method and instrument optimisation. This means that no one technique or method is suitable for the separation, identification and quantification of biomolecules related to a biological system of interest. To deal with this challenge, instrument methods are typically molecule-class specific such as amino acid-based, sugar-based, nucleic acid-based and lipid-based methods. Each of these methods involves specific steps to isolate, prepare and analyse these molecules.

As instrument method development is biomolecule class-specific, recent efforts have focused on comprehensive coverage of all class-specific biomolecules in the context of biology. The first of these efforts was the development of genomics, studying the complete set of DNA within a living organism. This field has been greatly accelerated as the result of these nucleic-acid specific instruments with DNA sequencing capable of determining the sequence of DNA with lengths up to 882,000 nucleotide bases long with 99.88% accuracy over 6.4 million nucleotide bases[8]. While genomics was first in the –omics field, the field of proteomics was established soon after.

The term proteomics was describe by Marc Wilkins in 1994 to describe a genomics-style analysis of proteins as part of an organism[9]. Before 1994, automated Edman degradation was the predominant method for peptide sequencing which encountered difficulties with modified amino-acids[10]. Since then, protein sequencing by mass-spectrometry has been a successor to Edman sequencing due in part to its greater sensitivity, robustness regarding modified amino acids and, in combination with a database of expected protein sequences, does not require sequencing of the entire protein for identification to be made. Further advances in mass-spectrometry instrumentation and bioinformatics have resulted in the diversification of mass-spectrometry applications in the field of proteomics with each subfield having their own specific instrument types and methods but the most dominant application of mass-spectrometry in the field of proteomics is currently bottom-up mass-spectrometric analysis[11].

Bottom-up mass-spectrometric peptide analysis typically involves enzymatic digestion of the protein into peptides with subsequent analysis using two forms of instrumentation, C18-based liquid chromatographic separation of the peptides and mass-spectrometry-based peptide detection and characterisation, with each subject to their own forms of optimisation[11]. The use of reversed-phase C18 HPLC allows separation of peptides based on hydrophobicity, reducing complexity before detection while the mass spectrometer detects ionised peptides followed by acquisition of tandem mass spectra and database searching to assign the most probably sequence. The resulting output is a chromatogram of peptide ions (to be used for identification and quantitation) as well as a series of fragmentation events of peptide ions detected (to be used for sequencing the peptides quantified).

The subsequent analysis of this data and transformation into biologically meaningful results has become almost entirely automated in proteomics, increasing the throughput of bottom-up analyses. The development of peptide search engines such as SequestHT[12] and Mascot[13], allows peptide sequence matching from a prepared database of expected proteins to fragmentation events in a chromatogram. Following this, these identifications are mapped to detected precursor peptide ions allowing the identifications to be assigned to quantified peaks across the chromatogram. Subsequently, these peptides can then be mapped to their corresponding protein sequences and protein abundance changes to be inferred. These abundance changes can then be mapped to protein

networks, to further understand the mechanism of the perturbations involved in the biological system of interest.

Similar to that described for proteomics, glycomics can also utilise similar mass-spectrometric techniques to achieve glycan detection and characterisation but there are a diverse range of other techniques with no clearly superior method. These analytical techniques are described in our published review (Paper 1[14]). Despite this, mass spectrometric analysis of glycans offers several benefits including simultaneous characterisation and quantification within a single run, much like the method used for bottom-up proteomics. Glycomics however, has its own challenges due to the vast majority of glycan structures being non-linear owing to the multiple carbohydrate linkages possible.

1.3 Automation and bioinformatics for MS-based glycomics

Despite the similarity with bottom-up protein analysis, there is a severe gap in the analysis of the glycomics data generated by LC-MS. While there are currently no databases that can be used for automated matching of MS/MS spectra based on *in-silico* fragmentation similar to the databases used for bottom-up proteomics, initiatives such as UniCarb-DB[15] are being developed to store annotated MS/MS spectra reported in glycomics publications. One challenge with an initiative such as this is the required reporting of essential metadata which influences the MS/MS spectra observed for a given glycan structure.

The MIRAGE initiative acknowledges the importance of metadata through encouraging the reporting of sample preparation and MS methods used to acquire the data[16, 17]. The MS reporting guideline covers all parameters that influence glycomic data acquisition including hardware used, the individual parameters for the software, how spectra were generated and the data interpreted[17]. Since 2017, five experimental research papers have cited these guidelines[18–22] but the method of reporting is largely manual and therefore inconsistent. Further development of MIRAGE and initiatives such as UniCarb-DB could help further develop automated spectral matching approaches for MS-based glycan analysis.

Even with the differences between proteomics and glycomics, and their associated methods, both fields have the same underlying goal: to answer biological questions through biomolecule detection, characterisation and quantification. The following paper evaluates current research methods and results for studying the role of protein-glycosylation in human diseases, and the recent technological development that has assisted these findings. In addition, we aimed to identify current areas of research which, if successful, can further strengthen the linking of glycomics to systems-oriented human disease research.

1.4 Paper 1 – Human disease glycomics: technology advances enabling protein glycosylation analysis – part 1

In paper 1, I wrote section 5 (Recent Advances in Glycan MS/MS Analysis), prepared Figure 1 and Table 2 and helped edit the manuscript. The overall manuscript design, planning and implementation were discussed by all authors. Overall responsibility was taken by first author, Arun Everest-Dass.

EXPERT REVIEW OF PROTEOMICS, 2018
VOL. 15, NO. 2, 165–182
https://doi.org/10.1080/14789450.2018.1421946

Taylor & Francis
Taylor & Francis Group

REVIEW

Human disease glycomics: technology advances enabling protein glycosylation analysis – part 1

Arun V Everest-Dass[a,b,c], Edward S X Moh[a,c], Christopher Ashwood[a,c], Abdulrahman M M Shathili[a,c] and Nicolle H Packer[a,b,c]

[a]Biomolecular Discovery and Design Research Centre, Faculty of Science and Engineering, Macquarie University, Sydney, Australia; [b]Institute for Glycomics, Griffith University, Gold Coast, Australia; [c]ARC Centre for Nanoscale BioPhotonics, Macquarie University, Sydney, Australia

ABSTRACT

Introduction: Protein glycosylation is recognized as an important post-translational modification, with specific substructures having significant effects on protein folding, conformation, distribution, stability and activity. However, due to the structural complexity of glycans, elucidating glycan structure-function relationships is demanding. The fine detail of glycan structures attached to proteins (including sequence, branching, linkage and anomericity) is still best analysed after the glycans are released from the purified or mixture of glycoproteins (glycomics). The technologies currently available for glycomics are becoming streamlined and standardized and many features of protein glycosylation can now be determined using instruments available in most protein analytical laboratories.

Areas covered: This review focuses on the current glycomics technologies being commonly used for the analysis of the microheterogeneity of monosaccharide composition, sequence, branching and linkage of released *N*- and *O*-linked glycans that enable the determination of precise glycan structural determinants presented on secreted proteins and on the surface of all cells.

Expert commentary: Several emerging advances in these technologies enabling glycomics analysis are discussed. The technological and bioinformatics requirements to be able to accurately assign these precise glycan features at biological levels in a disease context are assessed.

ARTICLE HISTORY
Received 29 October 2017
Accepted 22 December 2017

KEYWORDS
Glycomics; glycan structure; glycan analysis; mass spectrometry; disease glycomics

1. Glycosylation

All cells in nature are coated with glycans that are essential for biological processes and communication. These glycans are mainly distributed as a large array of glycoconjugates covalently bound to proteins and lipids, which include glycoproteins, proteoglycans, and glycolipids [1,2].

The thick layer of glycoconjugates on the cell surface is commonly referred to as the glycocalyx. These extensive, complex glycan structures were classically thought to only provide structural roles, but now it is understood that glycans participate in fundamental inter- and intracellular functions, such as protein quality control, adhesion, motility, endocytosis, and signal transduction [1,3]. Glycans are also known to affect cellular processes important in development, cell proliferation, differentiation, and morphogenesis [2,4]. Glycosylation changes are often associated with disease states, for example, cancer cells frequently display differently expressed glycans compared to those from normal cells [5]. Additionally, the evolved glycan differences between species and phylogeny are important markers for the immune system in discriminating between self and nonself [6].

Monosaccharides, the basic building blocks of glycans are linked in either linear or branched forms by glycosidic bonds [2]. Several sequential enzymatic actions of glycosidases and glycosyltransferases assemble diverse glycan structures from these basic units [7]. They can be further modified by chemical substitutions such as phosphorylation, sulfation, and acetylation. Mammalian glycans are assembled from only 10 core monosaccharides: fucose (Fuc), galactose (Gal), glucose (Glc), *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc), glucuronic acid (GlcA), iduronic acid (IdoA), mannose (Man), sialic acid (SA), and xylose (Xyl) . Their assembly is aided through an estimated action of 700 proteins involved in the glycosylation machinery to generate the full diversity of mammalian glycans (estimated to be ~7,000 protein-attached structures) [8,9]. This non-template-driven process produces heterogeneous glycan structures differing in sequence, branching, linkage, and anomericity.

Eukaryotic protein glycosylation is usually via *N*- and *O*-linkages. *N*-glycans are formed by the covalent linkage of oligosaccharide chains to the asparagine amino acid residue of the protein, commonly involving a GlcNAc residue and the consensus peptide sequence Asn-X-Ser/Thr (where X is any amino acid except proline). The *N*-linked glycans biosynthesis is initiated when the nascent protein enters the endoplasmic reticulum (ER). The *N*-glycans are initially synthesized as a lipid-linked oligosaccharide (LLO) precursor, and the glycans are transferred from LLO to a nascent polypeptide chain during translation and the block of sugars ($Glc_3Man_9GlcNAc_2$) is transferred to the amino group in the asparagine side chain. A series of processing reactions further trims the *N*-glycan in the ER. Interestingly, these initial steps are highly conserved across all eukaryotes [10]. The initial trimming of the $Glc_3Man_9GlcNAc_2$ moiety is through the removal of the Glc residues by glucosidases that reside in the lumen of the ER. Most

CONTACT Nicolle H Packer nicki.packer@mq.edu.au Department of Molecular Sciences, Macquarie University, North Ryde NSW-2109, Australia

glycoproteins exiting the ER carry nine or eight Man residues, depending on the action of the ER α-mannosidase I. In multicellular organisms, α1–2 mannosidases IA, IB, and 1C in the *cis*-Golgi further trim the *N*-glycan to give $Man_5GlcNAc_2$. The multitude of glycans ($Man_{5-9}GlcNAc_2$) through these trimming steps generate the so-called oligomannose-type *N*-glycans. Further processing of the $Man_5GlcNAc_2$ gives rise to the complex and hybrid types in the medial-Golgi by the action of an *N*-acetylglucosaminyltransferase (GlcNAcT-I), which adds a GlcNAc residue to the Man α1–3 in the core of $Man_5GlcNAc_2$. The addition of this residue enables the action of α-mannosidase II to remove the terminal α1–3 and α1–6 Man residues. Hybrid-type *N*-glycans are formed due to the incomplete action of this α-mannosidase II. Complex-type *N*-glycans initiated by antennae or branches by the addition of terminal GlcNAc residues by various *N*-acetylglucosaminyltranferases in the Golgi can create a plethora of branched structures. Further capping (Gal, Fuc, SAs) or modification of the core glycan (Fuc) occurs in the *trans*-Golgi. Thus the *N*-glycan heterogeneity arises from the combinatorial enzyme processing in the Golgi apparatus. For more detailed description of the glycosyltransferases, glycosidases, and the biosynthetic pathway, readers are encouraged to access the freely accessible online resource 'Essentials of Glycobiology' [10].

An *O*-glycan is commonly linked to the polypeptide mainly through a GalNAc residue (commonly referred to as mucin-type) to the hydroxyl group of serine or threonine amino acid residue [11]. Their biosynthesis follows protein *N*-glycosylation and folding. Unlike *N*-glycans, the synthesis of *O*-glycans is through stepwise addition of single monosaccharide residues. The initial step is the addition of the GalNac residue which is catalyzed by a polypeptide-*N*-acetyl-galactosaminyltransferase (ppGalNAcTs). There have been about 20 different ppGalNAcTs reported to initiate this process with varying tissue expression and differential specificities [12]. The GalNAc residues are further extended generating eight different cores; cores 1–4 are the most common in humans, while cores 5–8 have extremely restricted occurrences [13]. The *N*- and *O*-glycan cores are usually further extended into a variety of terminal glycans by the addition and substitution of various monosaccharides such as GlcNAc, Gal, GalNAc, Fuc, and SA. Sulfation of Gal and GlcNAc residues causes further diversification. The mature glycan now comprises of several biologically relevant determinants, such as the blood group antigens. They are often found to be the major regulation and recognition factor of the glycoconjugate [14].

Several other *O*-linked modifications of eukaryotic proteins have also been reported such as *O*-fucosylation, *O*-mannosylation, *O*-galactosylation, and *O*-GlcNAcylation which are distinct to mucin-type *O*-glycans. These other types of *O*-glycosylation are comparatively rare, require less building blocks, and are found on specific proteins. For example, *O*-GlcNAc is a dynamic protein modification found on specific cytoplasmic, nuclear, and mitochondrial proteins with key roles in cellular physiology and progression of many diseases [15]. Despite the importance of this modification, *O*-GlcNAc is difficult to detect by standard biochemical methods due to its lability in mass spectrometers [16].

The limited development of glycomics characterization compared to that of genomics and proteomics is usually attributed to the analytical challenges associated with the complete elucidation of diverse heterogeneous glycan structures. On the other hand, we are increasing our understanding of glycan-associated biological processes such as host–pathogen interactions and the immune response. Part 1 of this review summarizes methodological developments and technical innovations that have advanced our ability to structurally characterize glycans. While Part 2 gives examples of the application of these technologies to determine glycosylation changes occurring in a range of human diseases.

2. Detection of glycoconjugates

The various ways of detecting glycans on glycoconjugates include chemical reactions, metabolic labeling, and recognition by specific lectins or antibodies. Classic colorimetric chemical methods for detection of glycans on proteins include periodic acid–Schiff's (PAS) stain that requires the oxidation of vicinal hydroxyl groups on the monosaccharides; Alcian Blue, which is a charge-based stain typically used to visualize proteoglycans and mucopolysaccharides [17,18]; and the phenol and concentrated sulfuric acid (PSA) stain that is used to detect oligosaccharides, polysaccharides, and their derivatives at submicro amounts of sugars [19]. These colorimetric methods provide a quick look at the rough abundances of glycans (compared to glycoprotein standards) or distribution of carbohydrates across a histological preparation. However, detailed information of the specific glycoconjugate cannot be revealed using these methods alone.

Lectins and glycan-specific antibodies are used in glycan analysis to discriminate specific structural features of oligosaccharides such as α2-6-linked SA (*Sambucus nigra* lectin), sialyl-LewisX (E-selectin), and polysialic acids (mAb-735). These lectins (Table 1) and glycan-specific antibodies have been used in glycan visualization on histological slides, protein blots, and ELISA-type assays as a primary glycoconjugate-binding protein [20–22]. Broad specificity lectins are also used for enrichment of glycoproteins from a complex protein mixture for detailed analysis [23–25]. However, it is important to note that lectins and glycan-specific antibodies only react to the glyco-epitope that they target and do not necessarily reflect the whole glycoconjugate structure.

Glycan metabolic labeling using a chemical reporter such as an azide is another powerful tool to detect and track the glycosylation state of live cells [26] and animals [27,28] or specific glycoproteins [29]. Feeding nonnatural, acetylated monosaccharides containing the azide functional group (*N*-azidoacetylmannosamine, *N*-azidoacetylgalactosamine, *N*-azidoacetylglucosamine, and 6-azidofucose) triggers the cellular salvage pathway to convert them into azide-labeled nucleotide sugar substrates [30]. The glycan biosynthetic machinery then incorporates the azide-labeled sugars into the glycoconjugates and provides a unique functional group for specific targeting using strained-alkyne or phosphine chemical probes by biocompatible copper-free click chemistry. Apart from direct glycan imaging, glycan metabolic labeling also provides new tools for glycoprotein enrichment [31], glycoproteomics [32], and studies on *O*-GlcNAcylation [33].

Table 1. Commercially available lectins and their glycan structure specificity as described in the literature.

Lectin	Glycan affinity
Mannose-binding lectins	
Con A (Concanavalin A)	Branched α-mannosidic structures; high-mannose type, hybrid type and biantennary complex type *N*-?glycans
GNA (*Galanthus nivalis* lectin)	α1–3 and α1–6-linked high-mannose structures
LCH (Lentil lectin)	Fucosylated core region of bi- and triantennary complex type *N*-?glycans
Galactose/*N*-acetylgalactosamine-binding lectins	
RCA (*Ricinus communis* agglutinin)	Galβ1-4GlcNAcβ1-R
ECL (*Erythrina cristagalli* lectin)	Galβ1-4GlcNAcβ1-R
PNA (Peanut agglutinin)	Galβ1-3GalNAcα1-Ser/Thr (T-Antigen)
AIL (*Artocarpus integrifolia* lectin/ Jacalin)	(Neu5Ac)Galβ1-3GalNAcα1-Ser/Thr (T-Antigen)
VVL (*Vicia villosa* lectin)	GalNAcα-Ser/Thr (Tn-Antigen)
***N*-acetylglucosamine-binding lectins**	
WGA (Wheat germ agglutinin)	GlcNAcβ1-4GlcNAcβ1-4GlcNAc Neu5Ac
PHA (Phytohemagglutinin)	GlcNAc
Sialic acid - binding lectins	
SNA (*Sambucus nigra* lectin)	Neu5Acα2-6Gal(NAc)
MAL (*Maackia amurensis* leukoagglutinin)	Neu5Ac/Gcα2-3Galβ1-4GlcNAcβ1
Fucose-binding lectins	
UEA (*Ulex europaeus* agglutinin)	Fucα1-2Gal
AAL (*Aleuria aurantia* lectin)	Fucα1-2Galβ1-4(Fucα1-3/4)Galβ1-4GlcNAc GlcNAcβ1-4(Fucα1-6)GlcNAc

Use of azide-labeled nucleotide sugars has also been applied to therapeutic monoclonal antibodies as a novel method to create site-specific antibody-drug conjugates [34]. While glycan metabolic labeling provides great potential for visualizing the glycan landscape on cells, the underlying structural specifics of the glycome is not revealed and requires orthogonal methodologies for the investigation of the structural glycome.

3. Structural determination of protein glycosylation

The complexity associated with protein glycosylation analysis is due to the macro- and microheterogeneity of the oligosaccharides attached to the proteins. The microheterogeneity results from the range of glycan structures that can be attached at each glycosylation site on a protein, while the degree of site glycosylation confers the macroheterogeneity. The three usual approaches to study protein glycosylation are (a) characterization of intact glycoproteins, (b) characterization of protease-digested glycopeptides, and (c) structural analysis of glycans released from proteins.

3.1. Characterization of intact glycoproteins

Mass spectrometric (MS) characterization of intact glycoproteins is complicated by the extensive heterogeneity of the attached glycan moieties. The presence of these carbohydrate moieties further decreases the efficiency of ionization. The MS characterization of intact glycoproteins has been reported using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF), nanoelectrospray ionization (nESI), and capillary electrophoresis (CE) coupled to MS.

MALDI-based analysis of intact glycoproteins has been achieved with limited success for small proteins such as ribonuclease B, ovalbumin, bovine fetuin, and recombinant human erythropoietin [35–37]. These studies have revealed the significance of the proper matrix and optimum MS parameters that influence the detection of various glycoforms [36–38]. Novel approaches using in-source decay of intact glycoproteins induced by hydrogen radical transfer from the matrix produced several intense c-, y-, and z-type ions that enabled the site-specific assignment of mucin-type *O*-glycans on 17–18.5 kD MUC1 hexarepeat domains [39]. Analysis of large glycoproteins (~>100 kDa) with multiple glycosylation sites by MALDI-TOF-MS produces only an average molecular mass and is unable to resolve the molecular ions of the different glycoforms [40,41].

Glycoprotein analysis of intact glycoproteins by ESI-MS analyses is challenging due to the less-efficient desolvation resulting from the glycan heterogeneity and adduct formation in conventional ESI. The degree of glycosylation also limits detection, as glycans cover large surface areas of proteins thereby reducing efficient ionization. The resulting decreased charge state reduces the range of m/z analyzed by ESI-MS instruments. These technical difficulties are to some extent overcome by the application of nanoelectrospray as shown by Wilm and Mann [42] in the analysis of ovalbumin glycoforms. The coupling of nano-ESI with high-resolution mass analyzers such as time-of-flight (TOF) analyzers have also produced well-resolved glycoforms of bovine α1-acid glycoprotein [43] and cellulases purified from *Trichoderma reesi* [44]. Similarly, Nagy et al. showed the high-resolution α1-acid glycoprotein glycoforms by ESI-Fourier transform ion cyclotron resonance (FTICR) MS [45]. Heck and coworkers recently demonstrated the glycosylation analysis of native human erythropoietin using high-resolution native MS. The work elegantly demonstrated the characterization of site-specific glycans with minimal sample preparation and analysis time required to quantify glycan composition without ionization bias [46].

CE-MS at present provides great success for resolving the glycoforms of highly glycosylated proteins. Several studies have shown the application of CE-MS to almost completely resolving the various glycoforms of biologically relevant glycoproteins such as human plasma antithrombin [47] and recombinant erythropoietin [48,49]. The high resolving power of CE in analyzing glycoforms has potential for high-throughput screening of recombinant glycoproteins. One limitation is that there is less information obtained about the structural features of the attached glycans by this type of analysis. However, the integration of CE-MS with other orthogonal methodologies can mitigate this issue. For example, Takur et al. [50] demonstrated the characterization of 60 glycoforms of recombinant human chorionic gonadotrophin using CE coupled to a high-resolution FTICR MS; subsequent analysis of the tryptic glycopeptides enabled site-specific glycan variant identification.

3.2. Characterization of intact glycopeptides

Glycoproteins can be enzymatically digested to generate peptides, which overcomes the problem of mass-limited resolution of intact glycoprotein analysis in MS methods. Intact glycopeptide analysis (glycoproteomics) has the benefit of

allowing the identification of the site-specific heterogeneity of glycoproteins. There are several approaches to study glycopeptides by MS; typically, the sample is digested with specific endoproteinases (trypsin, chymotrypsin, Asp-N, Glu-C, and Lys-C) and the resulting peptides are analyzed by tandem MS [51]. Thaysen-Andersen et al. have recently reviewed in detail the advances in LC-MS/MS-based glycoproteomics [52], and will not be discussed in detail here. Similar to intact glycoprotein analysis by MS, intact glycopeptide analysis has its own challenges. Glycopeptides are often of low abundance compared to the non-glycosylated peptides and are represented with lower signal in the MS spectrum of the total peptide digest due to site-specific heterogeneity occurring at specific sites results in the signal intensities of the glycopeptides being relatively low. For example, human erythrocyte CD59 has over 100 different glycans on a single glycosylation site [53]. Furthermore, glycopeptides ionize poorly relative to the non-glycosylated peptides [54].

To alleviate these shortcomings, glycopeptides are usually enriched prior or during MS analysis [55]. Lectin affinity chromatography is a commonly used tool to remove non-glycosylated peptides and concentrate on the retained glycopeptides [24,56]. Typically, broadly selective lectins such as Con A are used [57,58] or alternatively, sequential or multi-lectin columns are used for the selective enrichment of glycopeptides [59,60]. Hydrophilic interaction chromatography (HILIC) is gaining importance in the enrichment of glycopeptides and relies on the added hydrophilicity conferred to glycopeptides due to the presence of the attached glycans. HILIC materials such as cellulose, sepharose, silica, aminopropyl, and zwitterionic types have been used for glycopeptide enrichment prior to MS analysis [61–64]. Boronic acid-based stationary phases also possess the potential to enrich for glycopeptides. Boronic acid is able to bind reversibly to vicinal *cis*-diol OH groups that are abundantly present in glycans. Boronic acid has recently been developed as the functionalized group on the surface of polymer particles or magnetic beads for chromatographic enrichment of both *N*- and *O*-glycopeptides [65–68].

Franc et al. showed that site-specific heterogeneity can also be determined by off-line separation and characterization [69]. They demonstrated this by simple fractionation of the hinge-region *O*-glycopeptides from IgA1 by reversed-phase liquid chromatography using a microgradient device and identified the glycopeptides by MALDI-TOF/TOF. This allowed the identification of isomeric *O*-glycoforms having the same molecular mass, but a different glycosylation pattern.

Recent developments in commercial availability of hybrid mass spectrometers with high-resolution mass analyzers have tremendously improved glycopeptide detection and analysis [70–72]. Researchers have attempted to use several fragmentation options such as higher-energy collision dissociation (HCD)/collision-induced dissociation (CID) and ETD sequentially to combine fragmentation information to decipher glycan- and peptide-specific information [73,74]. A data-dependent product ion trigger emanating from the high abundant oxonium ions, commonly observed in HCD fragmentation of glycopeptides, is used to activate subsequent ETD events [75]. This approach dedicates more instrument time for detailed ETD analysis of glycopeptides. The Heck lab recently introduced a hybrid dissociation method that combines ETD and HCD, called EThcD. After an initial electron-transfer dissociation step, all ions including the unreacted precursor ions are subjected to a supplemental energy which yields b/y- and c/z-type fragment ions in a single spectrum [76,77]. This unique hybrid fragmentation has already shown immense promise in glycopeptide analysis [72,78]. Although these novel methods provide valuable insight into site-specific occupation of glycan structures, very little information is currently available on the glycan structural features other than glycan composition. The glycan fragment ions for terminal substructures such as sialic acids or Lewis antigens can be detected but without any linkage information.

A proteomic strategy used to enrich and identify the sites of glycosylation on glycopeptides is by the selective covalent attachment of the glycopeptides to hydrazide resin after periodate oxidation of the carbohydrate moieties [79]. The glycosylation site is identified by releasing the attached *N*-glycans with the enzyme PNGaseF, and using MS to identify the deglycosylated peptides by the increased mass shift of the modified aspargine of +1Da in water or +3Da in $H_2{}^{18}O$. However, this approach is restricted to *N*-linked glycopeptide site identification and any information on the previously attached glycan moiety is lost. A related approach for identifying *O*-glycosylation sites is to release the glycans from the resin-bound *O*-linked glycopeptides using ammonia or ethylamine after PNGaseF release of the *N*-glycans [80,81]. The peptide integrity is maintained and the *O*-glycosylated serine or threonine site is labeled with a mass increment of 1Da (for ammonia) and 27Da (for ethylamine), that can be deduced by MS analysis of the deglycosylated peptides.

Currently the information garnered from intact glycopeptide analysis yields limited information on the fine structure of the attached glycan moiety, providing monosaccharide compositions but lacking detailed characterization of the terminal structural motifs. Nevertheless, qualitative and quantitative information on both the micro-and macroheterogeneity of glycoproteins can be obtained.

3.3. *Structural analysis of glycans released from proteins*

The characterization of *N*- and *O*-glycans released from proteins (glycomics) still remains the best currently available approach for the determination of the complete detailed glycan structural heterogeneity including monosaccharide composition, sequence, branching, and linkages as displayed on the surface of a cell. The remainder of this review (Part 1) will thus focus on the state-of-the-art and recent advances made in the characterization of the glycan structural motifs presented on the glycans released from proteins, and then give some examples in Part 2 of where this level of detail is required for an understanding of the difference between health and disease.

4. Characterization of the structure of glycan structural motifs

4.1. Release of *N*- and *O*-glycans from glycoproteins

The removal of *N*-linked glycans from proteins is facilitated by the use of the enzyme endoglycosidase peptide-*N*-glycosidase

F (PNGase F). This endoglycosidase cleaves *N*-glycans by hydrolysis of the glycosidic bond between the asparagine residue of a glycosylated protein and the reducing end GlcNAc of the attached oligosaccharide. *N*-glycans that contain an α1-3-linked Fuc attached to the reducing end GlcNAc (found in plants and insects) cannot be released by PNGase F but are released from glycopeptides by another endoglycosidase known as PNGase A [82]. There have been some recent advancements to this routine enzymatic method, with one approach employing ultra-high-pressure cycling, which subjects proteins to pressures of up to 30,000 psi and where glycoproteins become sufficiently denatured to be deglycosylated in as little as 20 min with PNGase F remaining active under these conditions [83]. Another method utilizes microwave radiation to assist the enzymatic cleavage of glycans from monoclonal antibodies in less than 10 min, whereas other glycoprotein standards required up to 1 h [84]. Recently, commercial companies such as New England Biolabs (Rapid PNGase F) and Bulldog Bio (PNGase F Prime) claim to have recombinantly engineered their enzymes for increased activity and wider coverage to enable rapid high-throughput analysis. A recombinant rice amidase called PNGase Ar (New England Biolabs) is able to remove glycans from intact plants glycoproteins and insect cells. Yan et al. showed that PNGase Ar was as effective in releasing *N*-glycans from *C. elegans* when compared to the use of hydrazine and far superior to traditional PNGase F and A release in this organism [85].

Chemical release methods are predominantly used to release the *O*-glycans from proteins, due to the lack of a general *O*-endoglycosidase. The only available endoglycosidase, *O*-glycanase, only cleaves the Core 1 type *O*-glycan (Galβ-3GalNAc) without any other modifications. Reductive β-elimination, using a strong alkaline reaction to cleave the glycoside-peptide bond, followed by reduction of the reducing terminus with $NaBH_4$ to prevent alkaline degradation of the released oligosaccharide, is the most widely used chemical method to release *O*-glycans [86]. An alternative approach is the glycan-releasing reaction of anhydrous hydrazine in which *N*- and/or *O*-glycans can be released under different conditions; *N*-glycans are released from glycoproteins using more vigorous conditions (85–100°C, 5–16 h) [87] whereas *O*-glycans are released and recovered under milder conditions (60°C, 6 h) [88]. Hydrazinolysis of *O*-glycans produces intact glycans with free reducing termini in high yield, but a major concern with this method is the occurrence of undesirable peeling of the reducing terminus and the restricted availability of the reagents. Other nonreductive chemical methods of release of *O*-glycans using dimethylamine [89] and ammonia [90] have also been described. These reactions claim to be quick, have low level of peeling, and provide free reducing ends that may be labeled.

The Cummings lab recently described a simple oxidative strategy using household bleach to release all types of free reducing *N*-glycans and *O*-glycan-acids from glycoproteins [91]. They showed by controlled treatment with sodium hypochlorite (NaClO) in commercial bleach, *N*- and *O*-glycans can be selectively released from glycoproteins. The further demonstrated the applicability of this method in releasing large quantities of glycans for chromatographic separation, MS sequencing and their use in glycan arrays for functional glycomics.

4.2. Nuclear magnetic resonance

Complete and unambiguous assignment of glycans including monosaccharide constituents, anomericity, linkages, and modifications is still best determined by nuclear magnetic resonance (NMR) [92]. In particular, NMR has gained significant importance in detecting relevant features in glycan-receptor interaction studies [93]. Recently, saturation transfer difference (STD)-NMR has been shown to be a robust and powerful approach to revealing the fine details of the molecular recognition between glycans and their binding receptors. For example, this technique provided valuable information for the design of inhibitors of Rhesus rotavirus and *Vibrio cholera* infections by enabling the detailed characterization of the molecular binding of various SA derivatives to Rotavirus VP8 subunit and *Vibrio cholera* sialidase, the causative agents of these two infections, respectively [94,95]. However, the major drawback of NMR analysis is the requirement of a large quantity of highly purified glycan, restricting its applicability as usually not enough sample can be obtained from biological sources [96].

4.3. Microarray-based analysis of glycans

Microarray-based technology is increasingly used in glycomics to identify glycan-binding partners and has been applied to the determination of glycan structural motifs. Glycan microarrays consist of immobilized glycans on glass slides and are used to characterize interactions between glycans and carbohydrate-binding proteins of interest [97,98,99]. Lectin and antibody-lectin sandwich microarrays have also been used to study glycosylation profiles. A typical lectin microarray consists of lectins immobilized as spots, either by activated aldehyde or epoxy groups, or adsorbed non-covalently on nitrocellulose and hydrogel slides. Glycoproteins, glycans, or whole cells are probed against the lectin microarray to identify specific glycan-binding structural motifs. Detection is usually carried out by fluorescent detection of the labeled ligands or through specific antibodies directed against the glycoprotein of interest [20,100]. Patwa et al. demonstrated the application of lectin microarray in the profiling of serum glycans [101]. Using evanescent-field fluorescence detection, weak lectin–glycan interactions under equilibrium conditions were measured at very low background, since the excitation wavelength of the evanescent field does not extend beyond 200 nm. Fluorescently labeled standard glycoproteins of known glycosylation were probed against 39 different lectins and analyzed quantitatively by this approach [102].

Lectin microarray profiling has multiple advantages over other glycan profiling techniques in terms of high throughput, low sample processing, and speed; however, it has also certain limitations, such as the low binding of lectins contributing to ligand loss during washing steps, low reproducibility, and limited glycan motif specificities. To the latter point, several studies have suggested that lectins have varying affinities toward several glycan motifs rather than absolute specificity for one structure [21,23].

4.4. Capillary electrophoresis

CE is a valuable separation system, widely used in the analysis of released glycans [103]. Commonly the glycans analyzed by CE are labeled by a charged chromophore or fluorophore such as 8-aminopyrene-1,3,6-trisulfonic acid, trisodium salt (APTS). A typical CE system consists of a silica separation capillary, a high voltage power supply, a detector, and an autosampler. The silica capillary contains surface silanol groups (Si-OH) that become ionized under electric potential. Usually, a very small amount of glycan sample analyte (1–50 nl) is required to separate well-resolved glycans in short run times [104]. Some drawbacks of this method for determining glycan structures include the need to pre-fractionate and label glycans and the high sample concentration required for detection post separation.

High isomeric selectivity can be obtained in CE analysis of fluorescently labeled glycans through the use of borate buffer that complexes with glycans. In noncomplexing buffers, the separation is usually attributed to size, charge, and to a lesser extent the shape of the glycan moiety [105]. Glycan peaks from CE are usually identified by (a) correlation of retention time with Glc units of a maltose ladder as reference; (b) comparison with glycan standards; (c) structure deduction through a series of exoglycosidase digestions; and (d) by coupling with detectors such as that used in DNA sequencers, laser-induced fluorescence (LIF) and MS [105].

4.5. Liquid chromatographic separation of glycans

High-performance liquid chromatography (HPLC) is an established separation technique for biomolecules and is frequently applied to the separation of glycans. Although HPLC has lower throughput capabilities than CE, it has a higher resolving power and better multiplexing with other analytical devices, such as easy front-end coupling with other chromatographic methods and back-end coupling to fluorescent detectors and MS. The various chromatographic techniques used to separate a complex mixture of underivatized glycans include high-pH anion exchange chromatography (HPAEC), size exclusion chromatography (SEC), HILIC, reversed phase (RP) and reversed-phase ion-pairing (RP-IP) chromatography, and porous graphitized carbon (PGC) chromatography. Similar to CE, an HPLC strategy can also be used for fluorescence-based quantitation and characterization of labeled released glycans. An internal dextran ladder standard combined with a bioinformatic platform such as GlycoBase/AutoGU [106] allows confirmation of glycan linkages and overall structure by retention time calibration and is available in a commercial package (Waters Corporation) [107].

Though most of these stand-alone separation methods have been extensively used to characterize and quantify glycans from complex mixtures using comparison with glycan standards and using retention time as the main parameter for identification, they have difficulty identifying novel or previously uncharacterized structures. Coupling these separation techniques to other detection methods such as MS mitigates this issue to an extent.

4.5.1. High-pH anion exchange chromatography

Similar to HPAEC-pulsed amperometric detection (PAD) analysis of monosaccharides, underivatized glycans can also be separated by HPAEC. The major drawback that arises through this type of analysis is the high amount of salt needed for elution and separation, which is incompatible with subsequent MS analysis. Although developments of desalting cartridges online prior to ESI-MS have been described for the analysis of *N*-glycans [108], this separation method is less favored. Maier et al. recently demonstrated the detection of low abundant *N*-glycans from various sources using a prototype 1 mm I.D. size HPAEC CarboPac PA200 column coupled to PAD detection or on-line ion trap (IT) MS for glycan structural assignment and quantification [109].

4.5.2. RP chromatography

Free glycans are too hydrophilic to be adsorbed by RP matrices. Reductive amination of glycans with a hydrophobic tag is used to allow the retention of glycans on RP columns. Several tags have been used, such as 2-aminopyridine (2-AP), 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA), and others [110,111]. The RP retention is due to the hydrophobicity of the tag, while resolution is based on the features of the glycan [112].

Permethylated *N*-glycans have also been separated by RP on a C18 column chip and analyzed by online MS [104,113]. This technique is promising, as permethylated glycans produce tandem mass spectra with high structural information. Some associated drawbacks are the under-permethylation of the glycans, possible desialylation of the acidic glycans and artifacts such as epimerization of the reducing end [114].

An additional strategy to retain glycans on RP is by ion pairing (IP). In RP-IP, charged additives in the eluting buffer are used as IP agents to increase the retention of oppositely charged analytes. RP-IP coupled to MS analysis has been used to characterize acidic glycans, especially negatively charged glycosaminoglycans [115,116].

4.5.3. Hydrophilic interaction chromatography

HILIC offers separation on the basis of glycan hydrophilicity, that is, dependent on properties such as size, charge, composition, structure, and linkage. HILIC is normal-phase chromatography, where an aqueous phase is usually used as the eluting solvent with an amide-derivatized silica as stationary phase.

Both charged and uncharged glycans can be separated by HILIC using a gradient from high to low organic content. The high reproducibility of retention times has allowed the high-throughput analysis of glycans using retention time libraries and standards. A high-throughput HILIC automated workflow in a robotic 96-well plate format with a workflow from sample processing through to data interpretation has been developed by Royle et al. [117]. These studies used fluorescently labeled glycans (2-aminobenzamide (2AB)) and detected them by fluorimetry. Wuhrer et al. showed that the structures of both labeled and underivatized glycans can be analyzed by the coupling of HILIC to ESI-MS [118].

There are several HILIC phases with interesting specificities and selectivities available for enriching hydrophilic molecules.

An interesting comparison test study was conducted by Ruhaak et al. to compare different HILIC phases (cellulose, sepharose, diol-bonded silica beads, and Biogel P-10) before subjecting the samples to a 48-channel multiplexed CGE-LIF analysis [119]. APTS-labeled glycans were purified from the reaction mixture that contained unused label, detergents, and proteins. Cellulose and Biogel P10 desalting gave the best results.

4.5.4. PGC chromatography

Porous glassy carbon as a chromatographic support matrix was modified by Knox and Gilbert to produce PGC that displayed good stability and chromatographic performance [120]. Earlier studies using PGC for chromatography suggested a reversed-phase-type behavior due to the proportional increase of retention of increasing hydrocarbons, but this does not explain the unique properties of isomeric and charge resolution. The retention mechanisms of glycans on PGC are still only vaguely comprehended with hydrophobic, ionic, and polar retention effects on graphite as known contributors. PGC shows superior resolution of native nonreduced and reduced glycans compared to conventional phases. Lately, the separation of permethylated glycans by PGC has also been reported [113,121]; other advancements include the packing of PGC into nanoscale chromatography chips for nano-LC-MS-based analysis of glycans [122].

The preparative sequential elution of neutral and acidic glycans using specific additives to the mobile phase was shown by Packer et al. [123], while Pabst and Altmann studied the influence of ionic strength, pH, and temperature on the retention of glycans by PGC [124]. They showed that when ionic strength is reduced while pH is maintained constant, acidic *N*-glycans are retained longer. Increased pH also increased retention time of acidic *N*-glycans with the neutral glycans unaffected, and a temperature increase resulted in stronger retention of all *N*-glycans. PGC has been shown to have remarkable selectivity in separating structurally similar glycan isomers. Several studies on both *N*- and *O*-glycans have described the baseline resolution of many structural glycan isomers [125,126].

An analytical development that has increased the sensitivity and improved reproducibility of nano-flow LC is the microfluidic chip [127]. The LC chip consists of a trapping column, switching valve, and LC column integrated and interfaced directly to the inlet of an ESI-based mass spectrometer. The chromatographic packing materials include reversed-phase, HILIC and PGC. These configurations have been used in several glycomic studies including the characterization of native milk oligosaccharides, permethylated glycans derived from blood serum glycoproteins from control individuals and late-stage breast cancer patients, and salivary glycoproteins [128]. A variant of this microfluidic chip incorporates an immobilized PNGase F reactor that readily releases glycans from intact IgG monoclonal antibodies in 6 s to claim a digestion efficiency of approximately 98% [129]. Unfortunately, this microfluidic chromatographic chip is no longer available from the manufacturer.

4.6. Glycosidase digestions

Structural elucidation of glycans by the enzymatic cleavage of specific exoglycosidases and endoglycosidases is a valuable strategy that complements CE, HPLC, and MS analyses [130]. Exoglycosidases are routinely used to confirm terminal glycan epitopes such as Lewis fucosylation linkages and SA linkages. An inherent disadvantage with this strategy is the multiple experimental analyses required, especially if sample has limited or low-abundance glycans.

4.7. MS analysis of released glycans

The coupling of MS with glycan separation methodologies has become the most favored and powerful technique for structural analysis of glycans. Some advantages of MS-based glycan analysis include high sensitivity, low sample requirement, and analytical versatility. MS analysis of peptides is relatively well defined and performed at high throughput, while, until reliable purpose-specific software is developed, glycan analysis by tandem MS data is particularly complex and labor-intensive to interpret. The tandem MS behavior is dependent on various factors such as product ion patterns, the type of derivative, and the fragmentation method.

The two main MS ionization techniques used currently for glycan analysis are MALDI and ESI.

4.7.1. Matrix-assisted laser desorption ionization

MALDI-MS-based structural characterization of glycans is an extensively used technique. Detailed information about the application of MALDI to glycan analysis, including matrices that are of particular use for carbohydrates, can be found in a series of comprehensive reviews by Harvey [131,132]. The MALDI-MS sample preparation is quite simple; the sample is mixed with a matrix solution and spotted onto a target plate. Crystals of the matrix-sample mixture are formed as the sample spots dry. Their strong optical absorption in either UV or IR range can rapidly and efficiently absorb the laser irradiation. A pulsed beam laser causes the matrix to transfer protons to the analyte molecules resulting in desorption and ionization of the sample molecules are extracted into the mass analyzer in the positive ionization mode. There are a wide variety of matrices available for glycan analysis, though the earliest developed matrix, 2,5-dihydroxybenzoic acid (DHB), remains the most popular. MALDI predominantly produces $[M+Na]^{1+}$ ions from positive mode ionization of glycans; other cations can also be introduced by doping the matrix with appropriate salts. Negative ions have also been reported, mainly from the use of β-carbolines as matrix.

A major advantage of MALDI is that contaminants and relatively high salt concentrations do not affect the sensitivity compared to other ionization methods. In addition, the glycan profile obtained is less complex because of the mainly singly charged ion formation. Coupling MALDI to a TOF/TOF analyzer can aid in accurate mass fragment analysis. Sialylated glycans however are quite labile when generated in the vacuum MALDI source without cooling of the newly formed ions, and have been reportedly lost in varying amounts in reflectron TOF, TOF/TOF, and qTof instruments [133,134]. To overcome

this problem, SAs can be permethylated, or the ions analyzed after they dissipate energy in an intermediate pressure MALDI source, prior to TOF analysis [135,136].

A recent development that promises to impact glycan analysis by MALDI-MS is the specific derivatization of the SA residues [137,138]. The reagent 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM) has been used in an esterification reaction with methanol that produces α2,3-linked SA lactones, whereas α2,6-linked SAs are esterified. This differential labeling creates a distinguishable 32 Da mass difference between α2,3- and α2,6-linked sialylated structures. Furthermore, these derivatized SAs were much more stable than their natural counterparts in MALDI analysis.

4.7.2. Electrospray ionization

Ionization by electrospray is obtained by introducing the liquid analyte by forced flow through a capillary with a fine nozzle. The high voltage applied to the nozzle creates a spray of charged droplets that enter the mass analyzer. These droplets are usually dried by high temperature, and by the flow of a drying gas, typically nitrogen. The electric charge density of the droplets increases as their size decreases, and eventually the mutual charge repulsion overcomes the surface tension (Rayleigh limit [43,139]) of the droplets resulting in their division [86].

ESI is usually referred to as a soft ionization technique, since it imparts relatively low energy to the analyte molecules during desolvation. Therefore, it is useful in the analysis of glycans with unstable modifications such as sialylation and sulfation [139]. Unlike MALDI-MS, ESI-MS is sensitive to salts and other contaminants, and therefore is usually best used after chromatography using volatile buffers. ESI usually produces multiply charged glycan ions that are advantageous for MS analyzers with limited mass range, albeit increasing the complexity of interpretation of the resultant spectra.

Depending on the ion source conditions and additives to solvents, ESI produces various types of ions. In the positive ion mode, both $[M + nH]^{n+}$ and their sodium adducts are formed under mild ion source conditions [140]. Under negative ion mode, glycans form $[M - nH]^{n-}$ ions or adducts with various anions. Harvey showed that nitrate adducts were particularly effective in the analysis of underivatized large *N*-glycans [141].

4.7.3. Mass analyzers

MS measurements are carried out in the gas phase on ionized analytes. The mass analyzer is used to separate the ions and is fundamental to the performance of the MS. The important features of mass analysis are sensitivity, accuracy, resolution, and the data acquisition speed. The various mass analyzers used in the analysis of glycans are IT, TOF, quadrupole, and FTICR analyzers. Each analyzer is different in design and capability, and can be used alone or in tandem with another.

Though IT have limited resolution, their MS^n fragmentation capabilities make them attractive analyzers when coupled to ESI. Modern Q-TOFs and Orbitraps possess higher resolution and their very high mass accuracy enables detailed characterization of glycans, for example, the phosphate and sulfate modifications could be distinguished natively without further sample processing [142]. Due to the increase in core facilities for proteomics and the desire to also perform glycan analysis, glycomics workflows that are compatible with instrument setups for proteomics have become increasingly popular. A strategy by Hsiao et al. [143] accomplishes delineation of permethylated glycan isomers using acidic RP C18 nanoLC separation. The use of high-resolution mass analyzers also allows for multiplexing of mass tags such as aminoxyTMT [144] and stable-isotope labeling [145].

5. Recent advances in glycan MS/MS analysis

The topic of tandem MS (MS/MS or molecular fragmentation) for glycan analysis has been recently reviewed [146] but in the context of glycan biomarkers of disease the potential of MS is the ability to validate not only the composition and sequence of all glycans in a complex mixture, but also to determine the branching and linkage structural glycoepitopes that are well-known to alter in many human diseases. This information is derived from tandem mass spectra of glycan analytes.

5.1. Fragmentation nomenclature

Given that fragmentation patterns of protein glycosylation can be quite complex, a systematic nomenclature of carbohydrate fragmentations was developed by Domon and Costello for FAB-MS/MS spectra and has since been adopted widely across the glycomics field [147]. Domon and Costello identified A, B, C, X, Y, and Z fragments; however, additional fragmentation pathways have since been characterized (Figure 1).

Spina et al. expanded upon fragmentation pathways defined by Domon and Costello to include a previously uncharacterized fragmentation mechanism in MALDI-TOF/TOF, denoting the fragments generated from this method as E, F, and G fragments which, in combination with existing A fragments, now allow identification of the linkage positions present and discrimination between isomeric sugars by MALDI [148]. Harvey et al. also identified D, E, and F fragments (denoted as D, E_2, F_2 for simplicity) as structurally informative features [149]. One such fragment, D, gives composition of the 6-antenna and identifies the presence of a bisecting GlcNAc.

The identification of these structurally important differentiating glycan structural fragments (examples in Table 2) has allowed glycobiologists to characterize glycan epitopes based on which fragments are present and to associate these detailed differences in glycan structure with many human diseases.

Despite the numerous structural fragments that can be generated from tandem MS of glycans, different fragmentation methods and glycan analyte forms can give varying fragment populations which impact on the level of structural detail that can be obtained from tandem MS.

Glycan fragmentation has been frequently studied in both CID and electron activated dissociation (ExD) modes. Each of these areas has had improvements in fragmentation with Orbitrap-based HCD and FTICR-based electron capture dissociation and electron detachment dissociation [151,152]. Another variable in fragmentation pattern and therefore structural information includes the sample preparation and sample

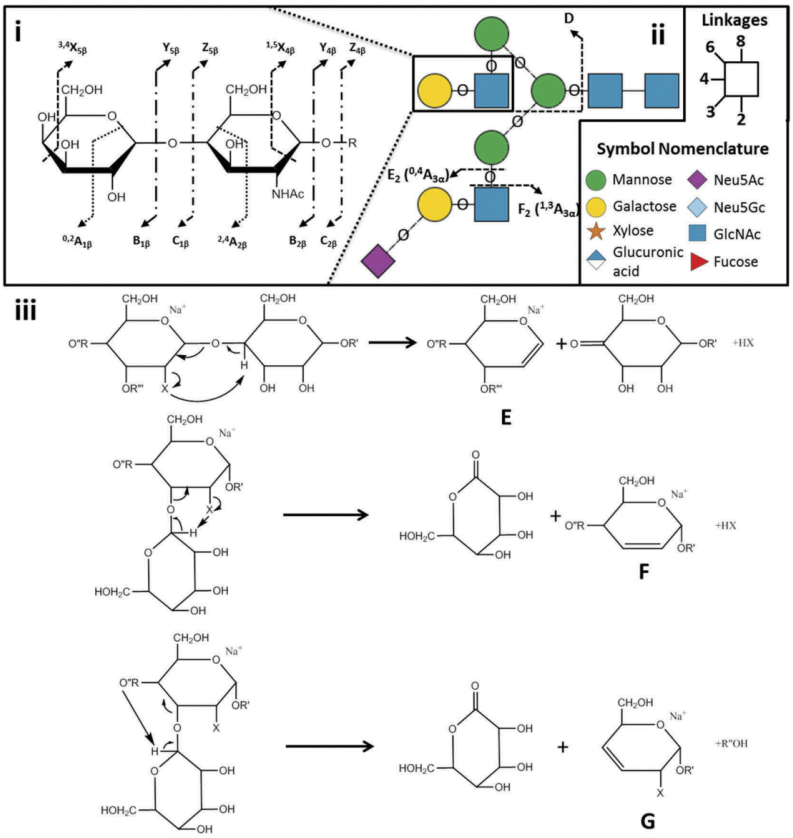


Figure 1. Fragmentation cleavages and nomenclature applicable to studying glycans. (i) Initial fragmentation nomenclature adapted from Domon and Costello [147]. (ii) Expansion of known carbohydrate fragments to include fragments indicative of antennae composition and glycan structure, adapted from Harvey et al. [149]. iii) Most recently identified fragmentation mechanisms involving gas phase rearrangement of glycan fragments, adapted from Spina et al. [148].

state such as permethylated glycans and reduction of the reducing end of native glycans. Both these variables impact the structural information obtained with tandem MS and have their own advantages and disadvantages.

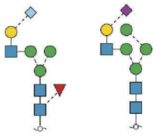
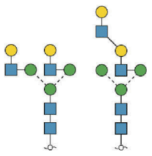
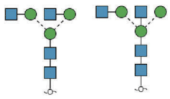
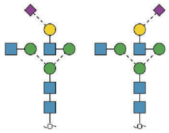
5.2. Collision-induced dissociation

CID (also known as collisionally activated dissociation) is the most commonly used fragmentation type for glycan analysis probably due to its availability across a wide range of mass analyzers provided by instrument manufacturers. This fragmentation method, as the name suggests, involves collision of molecular ions with neutral and nonreactive gases such as nitrogen, argon, and helium which generates structurally informative fragment ions. This technique results in the detection of all previously described fragments (A, B, C, D, E, F,G, X, Y, and Z) (Figure 1). The CID glycan spectra are dependent on the instrument type, and especially on the collisional energy [104] since the collision results in fragmentation of chemical bonds based on their lability. The resulting fragment ions are separated according to their m/z ratio and detected.

IT CID is one of the most frequently utilized tandem MS instruments used to characterize glycans partly due to its ability to perform MS^n, allowing glycans to be fragmented and structurally important ions to be further fragmented to confirm or ascertain more information. By using the MS in negative ion mode, more informative spectra can be generated through the generation of product ions consistent with glycosidic bond cleavage (B, C, Y, and Z fragments) and cross-ring cleavages (A and X fragments) [141]. Recent improvements in MS instruments containing IT have resulted in a wider variety of CID fragmentation methods such as being able to perform both beam-type and trap-type CID [153,154]. These improvements have resulted in a more versatile toolset for characterizing glycan structure and reporting of many different fragmentation patterns for even the same glycan.

Beam-type CID is a feature of triple quadrupole or QTOF-based instruments. These fragmentation techniques are

Table 2. Glycan structural information derived from tandem MS with examples of fragments that give this information.

Fragmentation information contribution	Example	Glycan-specific examples	Types of fragments	Example diagnostic ions
Composition	Identifying monosaccharide composition for isobaric glycans (both have same mass of 1730.6239 Da)		B, C, Y, and Z fragments	Neu5Ac (291 and 308 Da):  Neu5Gc (306 and 324 Da): 
Sequence	Presence of poly-*N*-acetyllactosamine chains		A, B, C, X, Y, and Z fragments	Poly-N-acetyllactosamine repeat (550 and 567 Da): 
Branching	Number of antennae and/or bisecting		A, B, C, X, Y, and Z fragments	D-221 ion (presence of bisecting GlcNAc (509 Da)) [149]: 
Linkages	α2,3 vs. α2,6 sialic acid linkage		A, B, C, E, F, G, X, Y, and Z fragments	Diagnostic ion for the presence of an α2,6 Neu5Ac isomer (307 Da) [150] 

denoted as being tandem-in-space due to the absence of a trapping mechanism which is commonly found in IT and FTICR mass analyzers which perform fragmentation tandem-in-time [155].

5.3. Electron activated dissociation

(ExD) of glycans contrasts with CID by providing fragmentation patterns based around cross-ring cleavages rather than glycosidic bond cleavage. ExD recently has been utilized to characterize and differentiate glycan isomers based on their linkage. In one example, CID and ExD were used in a complementary fashion to determine the composition, sequence, branching, and linkages of a permethylated disialylated biantennary *N*-glycan [152].

5.4. Standardization

Given the diverse types of glycomics analyses, the MIRAGE (Minimum Information Required for A Glycomics Experiment) initiative was developed to address the standardization of a wide range of glycosylation analysis challenges from sample preparation through to publication of data, including many aspects in between [156]. As a result of the glycomics community accepting these guidelines, it is hoped that the evaluation and reproduction of glycomics experiments will be improved as a result. Currently, three specific MIRAGE guidelines have been published on sample preparation [157], MS-based glycoanalytic data [158], and glycan microarrays [159].

5.5. Software tools

To complement the broad range of analytical techniques for quantitation of glycans, there is an accompanying variety of data analysis software being developed for analyzing data generated using these experimental techniques. For manual analysis of the data generated by these techniques, vendor-specific software is often used; however, as glycomics experiments become high throughput and more complicated, software packages dedicated to glycomics (varying from open-source to commercially available) are increasingly being developed. There are a few free software and databases available for glycomics, for example, ExPASy has provided a separate tab for glycomics (www.expasy.org/glycomics) that includes both internal and external links to various databases and tools, and a new website Glyconnect https://glyconnect.expasy.org/browser has been developed to integrate information on protein glycosylation structure and function from a range of bioinformatic resources. The UniCarb KnowledgeBase

(UniCarbKB) [160] is based on a curated database of glycan structures from glycoproteins as reported in scientific literature. An international collaborative effort has also been undertaken to provide a centralized resource for depositing glycan structures, compositions and topologies, and to retrieve accession numbers for each of these registered entries. This repository, called GlyTouCan [161,162], is available at http://glytoucan.org/. The emergence of these databases has been discussed in more detail elsewhere [163–166].

Software-assisted interpretation of glycan spectra can be performed by tools such as GlycoWorkBench [167], SimGlycan [168], Cartoonist [169], and MultiGlycan [170] among others [171] and UnicarbDB [172,173] collects annotated experimental MS/MS data on released glycans. Currently, there is little software available for high-throughput identification and quantitation of glycans and this remains an area that could be further improved; however, advances in proteomics software (Skyline [174] and MaxQuant [175]) could serve as useful starting platforms to apply to glycomics quantitation.

Although this review focuses predominantly on the structural characterization of glycans released from proteins, important intact glycopeptide analysis tools are briefly discussed here. GlycoMod [176], GlycoPep DB [177], and GlycoPep ID [178] are some of the freely accessible web-based tools for glycopeptide analysis. NetOGlyc and NetNGlyc are popular web-based tools for the prediction of potential glycan sites in the mammalian and human proteome [179,180]. GlycoMiner [181], Byonic [182], GlypID [183], and Protein Prospector [184] are promising software tools where sophisticated algorithms have been developed to characterize glycopeptides. Tsai et al. [171] have recently reviewed, in detail, the various features available in these and similar glycopeptide analysis software tools.

5.6. MALDI-MS tissue imaging of released glycans

Since its introduction, MS imaging (MSI) has provided unique advantages in the analysis of tissue specimens [185]. A broad spectrum of analytes ranging from proteins, peptides, glycans, lipids, small molecules, pharmaceutical compounds, endogenous, and exogenous metabolites can be analyzed *in situ* from tissue sections by this technique [186–194]. The data from such analysis is a pictograph that can be overlaid on the actual tissue sample and the location of the signal corresponds to the analyte's molecular mass. Although there are several types of MSI, the majority of analysis are performed using the MALDI-MSI approach. Any tissue section can be used for such analysis; typically these sections are covered with a suitable organic matrix material that co-crystallizes with the sample molecules. A laser beam irradiation ablates the matrix surface with the desorption and formation (for the most part) of singly protonated molecular species ($[M + H]^+$). These resulting ions traverse through a TOF analyzer, and their mass-to-charge (m/z) is determined. Using different matrices and technology (such as MALDI-TOF or MALDI-FT-ICR), the analyte classes can be chosen [195]. Typically a raster-scanned image of a tissue section provides a spatial resolution ranging from approximately 200 μm down to 20 μm generating a mass spectrum for every individually measured spot [196]. Recent technological improvements have contributed to imaging phospholipids, neuropeptides, and drug compounds at a pixel size between 5 and 10 µm [196–200].

MALDI-MSI of released *N*-glycans has been recently described by a few groups ([191,201–204]. Unique glycan m/z were clearly able to discriminate between cancer and noncancerous tissues [202,205,206]. MALDI-MSI was also used to image released glycan masses from Tissue MicroArrays (TMAs) of liver cancer in which many small tumor tissue samples of different patients were assembled on a single slide, demonstrating the high-throughput capability. On-tissue SA derivatization that differentiated between the linkage isomers of SA and also stabilized these acidic residues has also been successfully reported [207].

The correlation of MALDI MSI molecular information with other traditional histology of the same tissue section is a potential advantage offered by this technique. Furthermore, the label-free and multiplex analysis offers the simultaneous analysis of hundreds to thousands of molecules in a single analysis.

5.7. Ion-mobility MS

A rapidly developing MS-based technique is ion-mobility spectrometry (IMS). Through this setup, gas-phase ion analytes are subjected to a series of collisions under low-electric field, and the ions thereby separate due to their collisional cross sections and mass-to-charge values; a separation relative to the overall shape of the molecule. This technique in glycan analysis has found favorable application to isomer separation and characterization. One limitation of tandem-MS techniques is the dependence on separation of the isomers before fragmentation. This challenge appears to have been recently addressed by separation and identification of isomeric glycans by selected accumulation – trapped ion mobility MS (TIMS). In this case, a mixture of isomeric permethylated tetrasaccharides was separated using TIMS and characterized using ExD [208]. While the MS setup is largely experimental at this stage, it has potential for the discrimination of glycan isomers at the level of a single monosaccharide linkage difference. Other IMS studies clearly showed the advantage offered by the technique in the separation of branched oligosaccharides from linear oligosaccharides [209]. Three distinct permethylated Man_5 *N*-glycan isomers from ovalbumin were separated and reproducibly characterized by Plasencia et al. [210]. The group of Peter-Katalinic have used computer-assisted assignment of *N*- and *O*-free and amino acid-linked glycans. Unique patterns of glycans specific to their m/z values and drift times were observed and recorded by IMS-MS for the *de novo* identification of the human urinome [211]. Hinneburg et al. recently demonstrated separation of different *N*-acetylneuraminic acid linkages on *N*-glycans in a site-specific manner on individual glycopeptides using IMS analysis of diagnostic fragment ions [212].

A major advantage of IMS glycan analysis is its potential application to rapid clinical diagnosis. As early as 2012, Isailovic et al. used a combination of IMS and multivariate analysis with principal component analysis (PCA) to characterize serum *N*-linked glycans from 81 individuals (28 with cirrhosis of the liver, 25 with liver cancer, and 28 apparently healthy). Supervised PCA analysis of the combined ion-mobility profiles of different mass-to-charge ratios of glycan ions improved the

delineation of diseased states with an analysis time of 2 min per sample [213].

6. Expert commentary

The intrinsic complexity of glycan moieties has been a major factor in the progression of glycomics. The advancements in proteomics and genomics analysis were primarily due to the characterization of their respective linear codes by DNA sequencers and MS sequencing, but in glycomics, no single technique is yet capable of total and unbiased analysis.

The currently available technologies of glycan analysis require the release of glycans from their carrier proteins if the detailed structural glycan data in terms of composition, sequence, branching, and linkage are needed to be known. Determination of the heterogeneity of glycosylation at a site-specific level is currently reliant on glycopeptide analysis that provides important glycan composition and occupation data, but only partial glycan structural detail. The new derivatization methods on linkage-specific SAs (α2, 3 and α2, 6) are compatible with glycopeptides [214], and structural modifications on the *N*-glycan core such as core fucosylation and bisecting GlcNAc can now be identified from glycopeptide tandem MS fragmentation [51,215]; however, structural features such as antennae branching, linkage of outer arm fucosylation, Lewis epitopes, and polylactosamine extensions at particular protein sites in a complex mixture still require improvements in the analytical technologies currently available. The presence or addition of specific structural glycan features on proteins has been shown to be critical to many biological functions, but detailed analysis of the glycan heterogeneity present at a particular site is still required. Clearly, as far as can be seen, MS analysis is the core technology that will solve these shortcomings and we are seeing important steps in intact glycoprotein MS, ion mobility, and software development for spectral interpretation that promise to simplify the analysis of both the correct glycan structure and the specific site heterogeneity.

The recent years have brought about many inventive approaches and resourceful methods that enable the unraveling of the glycan code. To a large extent, glycan analysis complexities have been reduced and the role of specific glycan epitopes have been identified as a consequence.

7. Five-year view: a speculative viewpoint on how the field will evolve in 5 years time

In the field of proteomics, a more developed area of biomolecular MS, there have been recent developments in the availability of data through open-access platforms such as PRIDE. These platforms need to be developed so they can be used for glycomics analysis, providing open access to data and improving data sharing that will stimulate further method development in glycomics. Glycomics standardization initiatives such as MIRAGE are facilitating these developments; however, open-source and freely available informatics tools are required for this aim to be accomplished.

Analytical methodology for glycomic and glycoproteomic analyses will continue to improve. Since MS-based glycomic analyses require expensive equipment and a fair amount of expertise, other alternative technologies such as arrays and novel antibody-based assays will also continue to promote glycobiological knowledge. The MALDI-MSI application to glycan analysis is still in its primary stages of development, but shows enormous potential in determining the topological distribution of glycans in tissue. This is highly relevant especially in understanding tumor biology and for rapid classification of cancer types or subtypes.

The next major challenge is to assign function to the glycosylation differences that we are observing in every biological perturbation that we study. The problem is that there is no single answer and that every molecular interaction is being found to involve unique glycan structure(s), often at a particular site, density, or neighborhood. It is to be hoped that as more molecular biologists and medical researchers discover that glycosylation is involved in the fine-tuning of their particular biological system, there will be more emphasis on the determination of these precise molecular structures that modify proteins and that are integral to just about every biological system. And this review does not even address the challenges of analyzing glycan structures attached to other conjugates such as glycolipids, proteoglycans, peptidoglycans, and lipopolysaccharides!

Key issues

- Glycans are involved in cellular processes important in development, cell proliferation, differentiation and morphogenesis. The complexity associated with protein glycosylation analysis is due to the macro- and microheterogeneity of the oligosaccharides attached to the proteins.
- The limited development of glycomics compared to that of genomics and proteomics is usually attributed to the analytical challenges associated with the complete elucidation of diverse heterogeneous glycan structures.
- Mass spectrometric characterization of intact glycoproteins is complicated by the extensive heterogeneity of the attached glycan moieties. The presence of these carbohydrate moieties further decreases the efficiency of ionization.
- Similarly, glycopeptides also ionize poorly relative to the non-glycosylated peptides. To alleviate these shortcomings, glycopeptides are usually enriched prior or during mass spectrometry analysis.
- The characterization of N- and O- glycans released from proteins is currently the best approach to determine glycan structural heterogeneity as displayed on the surface of a cell, including monosaccharide composition, sequence, branching and linkages.
- High-performance liquid chromatography is an established separation technique applied routinely in the separation of glycans.
- The coupling of mass spectrometry with glycan separation methodologies has become the most favored and powerful technique for structural analysis of glycans. Some advantages of MS based glycan analysis include high sensitivity, low sample requirement and analytical versatility.

- Glycan mass spectrometry fragmentation spectra provide valuable information in the elucidation of their structure. Different fragmentation methods and glycan analyte forms can give varying fragment populations impacting the level of structural detail obtained.
- MALDI-MS imaging is a new promising technique that offers label-free multiplex analysis of several glycan species capable of delineating tissues by their spatial localisation.
- Ion mobility MS is another rapidly advancing technology that can be used to characterize and identify glycans from complex mixtures.
- Glycomics standardization initiatives are facilitating regularization of data reporting, however open-source, reliable and freely available informatics tools are required for this aim to be accomplished.
- Assigning function to the observed glycosylation differences is an enormous challenge that requires several orthogonal approaches and multi-disciplinary studies.

Funding

This manuscript was supported by a grant from the Australian Research Council Centre of Excellence in Nanoscale Biophotonics: CE140100003.

Declaration of interest

All authors have received funding from the Australian Research Council Centre of Excellence in Nanoscale Biophotonics and Macquarie University. A.V. Everest-Dass and N.H. Packer have additionally received funding from the Institute of Glycomics, Griffith University. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed. Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Brown JR, Crawford BE, Esko JD. Glycan antagonists and inhibitors: a fount for drug discovery. Crit Rev Biochem Mol Biol. 2007;42(6):481–515.
2. Varki A, Lowe JB. Biological roles of glycans. In: Varki A, Cummings RD, Esko JD, editors. Essentials of glycobiology. New York (NY): Cold Spring Harbor; 2009. p. 75–88.
3. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology. 1993;3(2):97–130.
•• **A comprehensive review on several glycan mediated roles. The review takes into account the several theories proposed for the different biological roles of glycans, and continues to explain that, while all of the theories are correct, exceptions to each can also be found.**
4. Hardy MR, Townsend RR, Lee YC. Monosaccharide analysis of glycoconjugates by anion exchange chromatography with pulsed amperometric detection. Anal Biochem. 1988;170(1):54–62.
5. Dube DH, Bertozzi CR. Glycans in cancer and inflammation–potential for therapeutics and diagnostics. Nat Rev Drug Discov. 2005;4(6):477–488.
6. Marth JD, Grewal PK. Mammalian glycosylation in immunity. Nat Rev Immunol. 2008;8(11):874–887.
7. Yarema KJ, Bertozzi CR. Characterizing glycosylation pathways. Genome Biol. 2001;2(5):REVIEWS0004.
8. Cummings RD. The repertoire of glycan determinants in the human glycome. Mol Biosyst. 2009;5(10):1087–1104.
9. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 2012;13(7):448–462.
10. Stanley P, Taniguchi N, Aebi M. N-glycans. In: Varki, A, Cummings, RD, editors. Essentials of glycobiology. 3rd ed. New York (NY):Cold Spring Harbor; 2015. p.99–111.
11. Schachter H. The joys of HexNAc. The synthesis and function of N- and O-glycan branches. Glycoconj J. 2000;17(7–9):465–483.
12. Brockhausen I, Schachter H, Stanley P, et al. O-GalNAc glycans. In: Varki A, Cummings RD, editors. Essentials of glycobiology. nd. New York (NY): Cold Spring Harbor; 2009.
13. Brockhausen I, Stanley P. O-GalNAc glycans. In: Varki A, Cummings RD, editors. Essentials of glycobiology. rd. New York (NY): Cold Spring Harbor; 2015. p. 113–123.
14. Varki A. Biological roles of glycans. Glycobiology. 2017;27(1):3–49.
15. Hart GW, Slawson C, Ramirez-Correa G, et al. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu Rev Biochem. 2011;80:825–858.
16. Zachara N, Akimoto Y, Hart GW, et al. The O-GlcNAc modification. In: Varki A, Cummings RD, editors. Essentials of glycobiology. rd. New York (NY): Cold Spring Harbor; 2015. p. 239–251.
17. Kilcoyne M, Gerlach JQ, Farrell MP, et al. Periodic acid-Schiff's reagent assay for carbohydrates in a microtiter plate format. Anal Biochem. 2011;416(1):18–26.
18. Dong W, Matsuno YK, Kameyama A. A procedure for Alcian blue staining of mucins on polyvinylidene difluoride membranes. Anal Chem. 2012;84(20):8461–8466.
19. Dubois M, Gilles K, Hamilton JK, et al. A colorimetric method for the determination of sugars. Nature. 1951;168(4265):167.
20. Ribeiro JP, Mahal LK. Dot by dot: analyzing the glycome using lectin microarrays. Curr Opin Chem Biol. 2013;17(5):827–831.
•• **A review focused on lectin microarray arrays and recent technological advancements.**
21. Gupta G, Surolia A, Sampathkumar SG. Lectin microarrays for glycomic analysis. Omics. 2010;14(4):419–436.
22. Sun Y, Cheng L, Gu Y, et al. A human lectin microarray for sperm surface glycosylation analysis. Mol Cell Proteomics. 2016;15(9):2839–2851.
23. Lee A, Nakano M, Hincapie M, et al. The lectin riddle: glycoproteins fractionated from complex mixtures have similar glycomic profiles. Omics. 2010;14(4):487–499.
24. Drake PM, Schilling B, Niles RK, et al. A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. Anal Biochem. 2011;408(1):71–85.
25. Lee LY, Hincapie M, Packer N, et al. An optimized approach for enrichment of glycoproteins from cell culture lysates using native multi-lectin affinity chromatography. J Sep Sci. 2012;35(18):2445–2452.
26. Baskin JM, Prescher JA, Laughlin ST, et al. Copper-free click chemistry for dynamic in vivo imaging. Proc Natl Acad Sci U S A. 2007;104(43):16793–16797.
27. Chang PV, Prescher JA, Sletten EM, et al. Copper-free click chemistry in living animals. Proc Natl Acad Sci U S A. 2010;107(5):1821–1826.
28. Laughlin ST, Baskin JM, Amacher SL, et al. In vivo imaging of membrane-associated glycans in developing zebrafish. Science. 2008;320(5876):664–667.
29. Belardi B, De La Zerda A, Spiciarich DR, et al. Imaging the glycosylation state of cell surface glycoproteins by two-photon fluorescence lifetime imaging microscopy. Angew Chem Int Ed Engl. 2013;52(52):14045–14049.
30. Laughlin ST, Bertozzi CR. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. Nat Protoc. 2007;2(11):2930–2944.
• **This detailed protocol illustrates the syntheses of the azido sugars *N*-azidoacetylmannosamine (ManNAz), *N*-azidoacetylgalactosamine (GalNAz), *N*-azidoacetylglucosamine (GlcNAz) and 6-**

azidofucose (6AzFuc), and the detection reagents phosphine-FLAG and phosphine-FLAG-His6 for cellular metabolic labeling and subsequent glycan-profiling and visualization.

31. Roper SM, Zemskova M, Neely BA, et al. Targeted glycoprotein enrichment and identification in stromal cell secretomes using azido sugar metabolic labeling. Proteomics Clin Appl. 2013;7(5–6):367–371.
32. Woo CM, Felix A, Zhang L, et al. Isotope-targeted glycoproteomics (IsoTaG) analysis of sialylated N- and O-glycopeptides on an Orbitrap Fusion Tribrid using azido and alkynyl sugars. Anal Bioanal Chem. 2017;409(2):579–588.
33. Palaniappan KK, Hangauer MJ, Smith TJ, et al. A chemical glycoproteomics platform reveals O-GlcNAcylation of mitochondrial voltage-dependent anion channel 2. Cell Rep. 2013;5(2):546–552.
34. Agarwal P, Bertozzi CR. Site-specific antibody-drug conjugates: the nexus of bioorthogonal chemistry, protein engineering, and drug development. Bioconjug Chem. 2015;26(2):176–192.
35. Sottani C, Fiorentino M, Minoia C. Matrix performance in matrix-assisted laser desorption/ionization for molecular weight determination in sialyl and non-sialyl oligosaccharide proteins. Rapid Commun Mass Spectrom. 1997;11(8):907–913.
36. Gimenez E, Benavente F, Barbosa J, et al. Towards a reliable molecular mass determination of intact glycoproteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom. 2007;21(16):2555–2563.
37. Gimenez E, Benavente F, Barbosa J, et al. Ionic liquid matrices for MALDI-TOF-MS analysis of intact glycoproteins. Anal Bioanal Chem. 2010;398(1):357–365.
38. Liu Z, Schey KL. Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis. J Am Soc Mass Spectrom. 2005;16(4):482–490.
39. Hanisch FG. Top-down sequencing of O-glycoproteins by in-source decay matrix-assisted laser desorption ionization mass spectrometry for glycosylation site analysis. Anal Chem. 2011;83(12):4829–4837.
40. Lee BS, Krishnanchettiar S, Lateef SS, et al. Characterization of oligosaccharide moieties of intact glycoproteins by microwave-assisted partial acid hydrolysis and mass spectrometry. Rapid Commun Mass Spectrom. 2005;19(18):2629–2635.
41. Sanz-Nebot V, Benavente F, Gimenez E, et al. Capillary electrophoresis and matrix-assisted laser desorption/ionization-time of flight-mass spectrometry for analysis of the novel erythropoiesis-stimulating protein (NESP). Electrophoresis. 2005;26(7–8):1451–1456.
42. Wilm M, Mann M. Analytical properties of the nanoelectrospray ion source. Anal Chem. 1996;68(1):1–8.
43. Karas M, Bahr U, Dulcks T. Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. Fresenius J Anal Chem. 2000;366(6–7):669–676.
44. Hui JP, White TC, Thibault P. Identification of glycan structure and glycosylation sites in cellobiohydrolase II and endoglucanases I and II from *Trichoderma reesei*. Glycobiology. 2002;12(12):837–849.
45. Nagy K, Vekey K, Imre T, et al. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry of human alpha-1-acid glycoprotein. Anal Chem. 2004;76(17):4998–5005.
46. Yang Y, Liu F, Franc V, et al. Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. Nat Commun. 2016;7:13397.

•• **This study provides an integrative approach, combining two advanced mass spectrometry-based methods, high-resolution native mass spectrometry, and middle-down proteomics, to analyze the micro-heterogeneity of intact glycoproteins such as human erythropoietin and the human plasma properdin.**

47. Demelbauer UM, Plematl A, Kremser L, et al. Characterization of glyco isoforms in plasma-derived human antithrombin by on-line capillary zone electrophoresis-electrospray ionization-quadrupole ion trap-mass spectrometry of the intact glycoproteins. Electrophoresis. 2004;25(13):2026–2032.
48. Balaguer E, Demelbauer U, Pelzing M, et al. Glycoform characterization of erythropoietin combining glycan and intact protein analysis by capillary electrophoresis - electrospray - time-of-flight mass spectrometry. Electrophoresis. 2006;27(13):2638–2650.
49. Neususs C, Demelbauer U, Pelzing M. Glycoform characterization of intact erythropoietin by capillary electrophoresis-electrospray-time of flight-mass spectrometry. Electrophoresis. 2005;26(7–8):1442–1450.
50. Thakur D, Rejtar T, Karger BL, et al. Profiling the glycoforms of the intact alpha subunit of recombinant human chorionic gonadotropin by high-resolution capillary electrophoresis-mass spectrometry. Anal Chem. 2009;81(21):8900–8907.
51. Wuhrer M, Catalina MI, Deelder AM, et al. Glycoproteomics based on tandem mass spectrometry of glycopeptides. J Chromatogr B Analyt Technol Biomed Life Sci. 2007;849(1–2):115–128.
52. Thaysen-Andersen M, Packer NH. Advances in LC-MS/MS-based glycoproteomics: getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. Biochim Biophys Acta. 2014;1844(9):1437–1452.

•• **This review focuses on the system-wide site-specific analysis of protein *N*- and *O*-linked glycosylation and the recent advances in LC-MS/MS-based glycoproteomics. A general discussion of experimental designs in glycoproteomics and sample preparation prior to LC-MS/MS is also covered.**

53. Rudd PM, Elliott T, Cresswell P, et al. Glycosylation and the immune system. Science. 2001;291(5512):2370–2376.
54. Toyama A, Nakagawa H, Matsuda K, et al. Deglycosylation and label-free quantitative LC-MALDI MS applied to efficient serum biomarker discovery of lung cancer. Proteome Sci. 2011;9:18.
55. Ito S, Hayama K, Hirabayashi J. Enrichment strategies for glycopeptides. Methods Mol Biol. 2009;534:195–203.
56. Kubota K, Sato Y, Suzuki Y, et al. Analysis of glycopeptides using lectin affinity chromatography with MALDI-TOF mass spectrometry. Anal Chem. 2008;80(10):3693–3698.
57. Catala C, Howe KJ, Hucko S, et al. Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. Proteomics. 2011;11(8):1530–1544.
58. Lee YC, Srajer Gajdosik M, Josic D, et al. Plasma membrane isolation using immobilized concanavalin A magnetic beads. Methods Mol Biol. 2012;909:29–41.
59. Qiu R, Regnier FE. Use of multidimensional lectin affinity chromatography in differential glycoproteomics. Anal Chem. 2005;77(9):2802–2809.
60. Durham M, Regnier FE. Targeted glycoproteomics: serial lectin affinity chromatography in the selection of O-glycosylation sites on proteins from the human blood proteome. J Chromatogr A. 2006;1132(1–2):165–173.
61. Wada Y, Tajiri M, Yoshida S. Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. Anal Chem. 2004;76(22):6560–6565.
62. Hagglund P, Bunkenborg J, Elortza F, et al. A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. J Proteome Res. 2004;3(3):556–566.
63. Calvano CD, Zambonin CG, Jensen ON. Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry. J Proteomics. 2008;71(3):304–317.
64. Mysling S, Palmisano G, Hojrup P, et al. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. Anal Chem. 2010;82(13):5598–5609.
65. Lee JH, Kim Y, Ha MY, et al. Immobilization of aminophenylboronic acid on magnetic beads for the direct determination of glycoproteins by matrix assisted laser desorption ionization mass spectrometry. J Am Soc Mass Spectrom. 2005;16(9):1456–1460.
66. Monzo A, Bonn GK, Guttman A. Boronic acid-lectin affinity chromatography. 1. Simultaneous glycoprotein binding with selective or combined elution. Anal Bioanal Chem. 2007;389(7–8):2097–2102.

67. Ozohanics O, Turiak L, Drahos L, et al. Comparison of glycopeptide/glycoprotein enrichment techniques. Rapid Commun Mass Spe. 2012;26(2):215–217.
68. Desaire H. Glycopeptide analysis: recent developments and applications. Mol Cell Proteomics. 2013 Apr;12(4):893-901.
69. Franc V, Rehulka P, Raus M, et al. Elucidating heterogeneity of IgA1 hinge-region O-glycosylation by use of MALDI-TOF/TOF mass spectrometry: role of cysteine alkylation during sample processing. J Proteomics. 2013;92:299–312.
70. Zhao P, Viner R, Teo CF, et al. Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. J Proteome Res. 2011;10(9):4088–4104.
71. Darula Z, Sherman J, Medzihradszky KF. How to dig deeper? Improved enrichment methods for mucin core-1 type glycopeptides. Mol Cell Proteomics. 2012;11(7):O111 016774.
72. Yu Q, Wang B, Chen Z, et al. Electron-transfer/higher-energy collision dissociation (EThcD)-enabled intact glycopeptide/glycoproteome characterization. J Am Soc Mass Spectrom. 2017;28(9):1751–1764.
73. Saba J, Dutta S, Hemenway E, et al. Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation. Int J Proteomics. 2012;560391:2012.
74. Alley WR Jr., Mechref Y, Novotny MV. Characterization of glycopeptides by combining collision-induced dissociation and electron-transfer dissociation mass spectrometry data. Rapid Commun Mass Spe. 2009;23(1):161–170.
75. Singh C, Zampronio CG, Creese AJ, et al. Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins. J Proteome Res. 2012;11(9):4517–4525.
76. Frese CK, Altelaar AF, Van Den Toorn H, et al. Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. Anal Chem. 2012;84(22):9668–9673.
77. Mommen GP, Frese CK, Meiring HD, et al. Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD). Proc Natl Acad Sci U S A. 2014;111(12):4507–4512.
78. Marino F, Bern M, Mommen GPM, et al. Extended O-GlcNAc on HLA Class-I-bound peptides. J Am Chem Soc. 2015;137(34):10922–10925.
79. Nilsson J, Ruetschi U, Halim A, et al. Enrichment of glycopeptides for glycan structure and attachment site identification. Nat Methods. 2009;6(11):809–811.
80. Hanisch FG, Jovanovic M, Peter-Katalinic J. Glycoprotein identification and localization of O-glycosylation sites by mass spectrometric analysis of deglycosylated/alkylaminylated peptide fragments. Anal Biochem. 2001;290(1):47–59.
81. Rademaker GJ, Pergantis SA, Blok-Tip L, et al. Mass spectrometric determination of the sites of O-glycan attachment with low picomolar sensitivity. Anal Biochem. 1998;257(2):149–160.
82. Tretter V, Altmann F, Marz L. Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1-3 to the asparagine-linked N-acetylglucosamine residue. Eur J Biochem/FEBS. 1991;199(3):647–652.
83. Szabo Z, Guttman A, Karger BL. Rapid release of N-linked glycans from glycoproteins by pressure-cycling technology. Anal Chem. 2010;82(6):2588–2593.
84. Zhou H, Briscoe AC, Froehlich JW, et al. PNGase F catalyzes de-N-glycosylation in a domestic microwave. Anal Biochem. 2012;427(1):33–35.
85. Yan S, Vanbeselaere J, Wols F, et al. Core richness of N-glycans of *Caenorhabditis elegans*: a case study on chemical and enzymatic release. Anal Chem. 2018 Jan 2;90(1):928-935.
86. Carlson DM. Oligosaccharides isolated from pig submaxillary mucin. J Biol Chem. 1966;241(12):2984–2986.
87. Patel T, Bruce J, Merry A, et al. Use of hydrazine to release in intact and unreduced form both N- and O-linked oligosaccharides from glycoproteins. Biochemistry. 1993;32(2):679–693.
88. Alley WR Jr., Madera M, Mechref Y, et al. Chip-based reversed-phase liquid chromatography-mass spectrometry of permethylated N-linked glycans: a potential methodology for cancer-biomarker discovery. Anal Chem. 2010;82(12):5095–5106.
89. Zauner G, Koeleman CA, Deelder AM, et al. Mass spectrometric O-glycan analysis after combined O-glycan release by beta-elimination and 1-phenyl-3-methyl-5-pyrazolone labeling. Biochim Biophys Acta. 2012;1820(9):1420–1428.
90. Wang C, Fan W, Zhang P, et al. One-pot nonreductive O-glycan release and labeling with 1-phenyl-3-methyl-5-pyrazolone followed by ESI-MS analysis. Proteomics. 2011;11(21):4229–4242.
91. Song X, Ju H, Lasanajak Y, et al. Oxidative release of natural glycans for functional glycomics. Nat Methods. 2016;13(6):528–534.
92. Battistel MD, Azurmendi HF, Yu B, et al. NMR of glycans: shedding new light on old problems. Prog Nucl Magn Reson Spectrosc. 2014;79:48–68.
93. Gimeno A, Reichardt NC, Canada FJ, et al. NMR and molecular recognition of N-glycans: remote modifications of the saccharide chain modulate binding features. ACS Chem Biol. 2017;12(4):1104–1112.
94. Haselhorst T, Wilson JC, Thomson RJ, et al. Saturation transfer difference (STD) 1H-NMR experiments and in silico docking experiments to probe the binding of N-acetylneuraminic acid and derivatives to *Vibrio cholerae* sialidase. Proteins. 2004;56(2):346–353.
95. Haselhorst T, Blanchard H, Frank M, et al. STD NMR spectroscopy and molecular modeling investigation of the binding of N-acetylneuraminic acid derivatives to rhesus rotavirus VP8* core. Glycobiology. 2007;17(1):68–81.
96. Mulloy B, Dell A, Stanley P, et al. Structural analysis of glycans. In: Varki A, Cummings RD, editors. Essentials of glycobiology. rd. New York (NY): Cold Spring Harbor; 2015. p. 639–652.
97. Oyelaran O, Gildersleeve JC. Glycan arrays: recent advances and future challenges. Curr Opin Chem Biol. 2009;13(4):406–413.
98. Liang CH, Wu CY. Glycan array: a powerful tool for glycomics studies. Expert Rev Proteomics. 2009;6(6):631–645.
99. Day CJ, Paton AW, Higgins MA, Shewell LK, Jen FE, Schulz BL, Herdman BP, Paton JC, and Jennings MP. Structure aided design of a Neu5Gc specific lectin. Sci rep. 2017;7(1):1495.28473713, PMID 28473713.
100. Feizi T, Fazio F, Chai W, et al. Carbohydrate microarrays - a new set of technologies at the frontiers of glycomics. Curr Opin Struct Biol. 2003;13(5):637–645.
101. Patwa TH, Zhao J, Anderson MA, et al. Screening of glycosylation patterns in serum using natural glycoprotein microarrays and multi-lectin fluorescence detection. Anal Chem. 2006;78(18):6411–6421.
102. Kuno A, Uchiyama N, Koseki-Kuno S, et al. Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. Nat Methods. 2005;2(11):851–856.
103. Zamfir A, Peter-Katalinic J. Capillary electrophoresis-mass spectrometry for glycoscreening in biomedical research. Electrophoresis. 2004;25(13):1949–1963.
104. Bielik A, Zaia J. Historical overview of glycoanalysis. In: Li J, editor. Functional glycomics. New York, USA: Humana Press; 2010. p. 9–30.
105. Zaia J. Capillary electrophoresis-mass spectrometry of carbohydrates. Methods Mol Biol. 2013;984:13–25.
106. Campbell MP, Royle L, Radcliffe CM, et al. GlycoBase and autoGU: tools for HPLC-based glycan analysis. Bioinformatics. 2008;24(9):1214–1216.
107. Lauber MA, Yu YQ, Brousmiche DW, et al. Rapid preparation of released N-glycans for HILIC analysis using a labeling reagent that facilitates sensitive fluorescence and ESI-MS detection. Anal Chem. 2015;87(10):5401–5409.
108. Guignard C, Jouve L, Bogeat-Triboulot MB, et al. Analysis of carbohydrates in plants by high-performance anion-exchange chromatography coupled with electrospray mass spectrometry. J Chromatogr A. 2005;1085(1):137–142.
109. Maier M, Reusch D, Bruggink C, et al. Applying mini-bore HPAEC-MS/MS for the characterization and quantification of Fc N-glycans

from heterogeneously glycosylated IgGs. J Chromatogr B Analyt Technol Biomed Life Sci. 2016;1033–1034:342–352.
110. Pabst M, Kolarich D, Poltl G, et al. Comparison of fluorescent labels for oligosaccharides and introduction of a new postlabeling purification method. Anal Biochem. 2009;384(2):263–273.
111. Marino K, Bones J, Kattla JJ, et al. A systematic approach to protein glycosylation analysis: a path through the maze. Nat Chem Biol. 2010;6(10):713–723.
112. Bleckmann C, Geyer H, Geyer R. Nanoelectrospray-MS(n) of native and permethylated glycans. Methods Mol Biol. 2011;790:71–85.
113. Huang Y, Zhou S, Zhu J, et al. LC-MS/MS isomeric profiling of permethylated N-glycans derived from serum haptoglobin of hepatocellular carcinoma (HCC) and cirrhotic patients. Electrophoresis. 2017 Sep;38(17):2160-2167.
114. Ashline DJ, Lapadula AJ, Liu YH, et al. Carbohydrate structural isomers analyzed by sequential mass spectrometry. Anal Chem. 2007;79(10):3830–3842.
115. Kuberan B, Lech M, Zhang L, et al. Analysis of heparan sulfate oligosaccharides with ion pair-reverse phase capillary high performance liquid chromatography-microelectrospray ionization time-of-flight mass spectrometry. J Am Chem Soc. 2002;124(29):8707–8718.
116. Thanawiroon C, Rice KG, Toida T, et al. Liquid chromatography/mass spectrometry sequencing approach for highly sulfated heparin-derived oligosaccharides. J Biol Chem. 2004;279(4):2608–2615.
117. Royle L, Campbell MP, Radcliffe CM, et al. HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. Anal Biochem. 2008;376(1):1–12.
118. Wuhrer M, Koeleman CAM, Hokke CH, et al. Nano-scale liquid chromatography-mass spectrometry of 2-aminobenzamide-labeled oligosaccharides at low femtomole sensitivity. Int J Mass Spectrom. 2004;232(1):51–57.
119. Reusch D, Haberger M, Kailich T, et al. High-throughput glycosylation analysis of therapeutic immunoglobulin G by capillary gel electrophoresis using a DNA analyzer. MAbs. 2014;6(1):185–196.
120. Gilbert MT, Knox JH, Kaur B. Porous glassy carbon, a new columns packing material for gas chromatography and high-performance liquid chromatography. Chromatographia. 1982;16(1):138–146.
121. Costello CE, Contado-Miller JM, Cipollo JF. A glycomics platform for the analysis of permethylated oligosaccharide alditols. J Am Soc Mass Spectrom. 2007;18(10):1799–1812.
122. Hua S, An HJ, Ozcan S, et al. Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. Analyst. 2011;136(18):3663–3671.
123. Packer NH, Lawson MA, Jardine DR, et al. A general approach to desalting oligosaccharides released from glycoproteins. Glycoconj J. 1998;15(8):737–747.
124. Pabst M, Altmann F. Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides. Anal Chem. 2008;80(19):7534–7542.
125. Jensen PH, Karlsson NG, Kolarich D, et al. Structural analysis of N- and O-glycans released from glycoproteins. Nat Protoc. 2012;7(7):1299–1310.
126. Pabst M, Bondili JS, Stadlmann J, et al. Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. Anal Chem. 2007;79(13):5051–5057.
127. Hua S, Lebrilla C, An HJ. Application of nano-LC-based glycomics towards biomarker discovery. Bioanalysis. 2011;3(22):2573–2585.
128. Ruhaak LR, Miyamoto S, Lebrilla CB. Developments in the identification of glycan biomarkers for the detection of cancer. Mol Cell Proteomics. 2013;12(4):846–855.
129. Bynum MA, Yin H, Felts K, et al. Characterization of IgG N-glycans employing a microfluidic chip that integrates glycan cleavage, sample purification, LC separation, and MS detection. Anal Chem. 2009;81(21):8818–8825.
130. Greenwell P. Blood group antigens: molecules seeking a function?. Glycoconj J. 1997;14(2):159–173.
131. Harvey DJ. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update for 2011–2012. Mass Spectrom Rev. 2017;36(3):255–422.
•• **A continuing widespread review about the analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry.**
132. Harvey DJ. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update for 2009–2010. Mass Spectrom Rev. 2015;34(3):268–422.
133. Nie H, Li Y, Sun XL. Recent advances in sialic acid-focused glycomics. J Proteomics. 2012;75(11):3098–3112.
134. O'Connor PB, Mirgorodskaya E, Costello CE. High pressure matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for minimization of ganglioside fragmentation. J Am Soc Mass Spectrom. 2002;13(4):402–407.
135. Harvey DJ. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update covering the period 1999–2000. Mass Spectrom Rev. 2006;25(4):595–662.
136. Zaia J. Mass spectrometry and the emerging field of glycomics. Chem Biol. 2008;15(9):881–892.
137. Reiding KR, Blank D, Kuijper DM, et al. High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. Anal Chem. 2014;86(12):5784–5793.
138. Bladergroen MR, Reiding KR, Hipgrave Ederveen AL, et al. Automation of high-throughput mass spectrometry-based plasma N-glycome analysis with linkage-specific sialic acid esterification. J Proteome Res. 2015;14(9):4080–4086.
139. Cech NB, Enke CG. Practical implications of some recent studies in electrospray ionization fundamentals. Mass Spectrom Rev. 2001;20(6):362–387.
140. Dell A. Preparation and desorption mass spectrometry of permethyl and peracetyl derivatives of oligosaccharides. Methods Enzymol. 1990;193:647–660.
141. Harvey DJ. Fragmentation of negative ions from carbohydrates: part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates. J Am Soc Mass Spectrom. 2005;16(5):622–630.
142. Bossio RE, Marshall AG. Baseline resolution of isobaric phosphorylated and sulfated peptides and nucleotides by electrospray ionization FTICR MS: another step toward mass spectrometry-based proteomics. Anal Chem. 2002;74(7):1674–1679.
143. Hsiao CT, Wang PW, Chang HC, et al. Advancing a high throughput glycotope-centric glycomics workflow based on nanoLC-MS(2)-product dependent-MS(3) analysis of permethylated glycans. Mol Cell Proteomics. 2017;16(12):2268–2280.
144. Zhou S, Hu Y, Veillon L, et al. Quantitative LC-MS/MS glycomic analysis of biological samples using aminoxyTMT. Anal Chem. 2016;88(15):7515–7522.
145. Hu Y, Desantos-Garcia JL, Mechref Y. Comparative glycomic profiling of isotopically permethylated N-glycans by liquid chromatography/electrospray ionization mass spectrometry. Rapid Commun Mass Spe. 2013;27(8):865–877.
146. Leymarie N, Zaia J. Effective use of mass spectrometry for glycan and glycopeptide structural analysis. Anal Chem. 2012;84(7):3040–3048.
147. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate Journal. 1988;5(4):397–409.
•• **Describes the fragmentation nomenclature for glycans.**
148. Spina L, Romeo DI, Garozzo G, et al. New fragmentation mechanisms in matrix-assisted laser desorption/ionization time-of-flight/time-of-flight tandem mass spectrometry of carbohydrates. *Rapid Communications in Mass Spectrometry*. 2004;18:392–398.
149. Harvey DJ, Royle L, Radcliffe CM, et al. Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. Anal Biochem. 2008;376:44–60.
•• **Review of *N*-glycan fragmentation features.**

150. Wheeler SF, Harvey DJ. Negative ion mass spectrometry of sialylated carbohydrates: discrimination of N-acetylneuraminic acid linkages by MALDI-TOF and ESI-TOF mass spectrometry. Anal Chem. 2000;72(20):5027–5039.
151. Hu Q, Noll RJ, Li H, et al. The Orbitrap: a new mass spectrometer. J Spectrom. 2005;40(4):430–443.
152. Yu X, Jiang Y, Chen Y, et al. Detailed glycan structural characterization by electronic excitation dissociation. Anal Chem. 2013;85(21):10017–10021.
153. Olsen JV, Schwartz J, Griep-Raming J, et al. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. Mol Cell Proteomics. 2009;8(12):2759–2769.
154. Löner C, Blackstock W, Gunaratne J. Enhanced performance of pulsed q collision induced dissociation-based peptide identification on a dual-pressure linear ion trap. J Am Soc Mass Spectrom. 2012;23(1):186–189.
155. Johnson JV, Yost R, Kelley PE, et al. Tandem-in-space and tandem-in-time mass spectrometry: triple quadrupoles and quadrupole ion traps. Anal Chem. 1990;62(20):2162–2172.
156. York WS, Agravat S, Aoki-Kinoshita KF, et al. MIRAGE: the minimum information required for a glycomics experiment. Glycobiology. 2014;24(5):402–406.
157. Struwe WB, Agravat S, Aoki-Kinoshita KF, et al. The minimum information required for a glycomics experiment (MIRAGE) project: sample preparation guidelines for reliable reporting of glycomics datasets. Glycobiology. 2016;26(9):907–910.
158. Kolarich D, Rapp E, Struwe WB, et al. The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data. Mol Cell Proteomics. 2013;12:991–995.
•• **Glycan standardization initiatives to improving the standards for reporting MS-based glycoanalytic data.**
159. Liu Y, McBride R, Stoll M, et al. The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting glycan microarray-based data. Glycobiology. 2016;27(4):280–284.
160. Campbell MP, Peterson R, Mariethoz J, et al. UniCarbKB: building a knowledge platform for glycoproteomics. Nucleic Acids Res. 2014;42(Database issue):D215–221.
161. Aoki-Kinoshita K, Agravat S, Aoki NP, et al. GlyTouCan 1.0–the international glycan structure repository. Nucleic Acids Res. 2016;44(D1):D1237–1242.
162. Tiemeyer M, Aoki K, Paulson J, et al. GlyTouCan: an accessible glycan structure repository. Glycobiology. 2017;27(10):915–919.
163. Campbell MP, Peterson RA, Gasteiger E, et al. Navigating the glycome space and connecting the glycoproteome. Methods Mol Biol. 2017;1558:139–158.
164. Lisacek F, Mariethoz J, Alocci D, et al. Databases and associated tools for glycomics and glycoproteomics. Methods Mol Biol. 2017;1503:235–264.
165. Barnett CB, Aoki-Kinoshita KF, Naidoo KJ. The glycome analytics platform: an integrative framework for glycobioinformatics. Bioinformatics. 2016;32(19):3005–3011.
166. Lutteke T. Handling and conversion of carbohydrate sequence formats and monosaccharide notation. Methods Mol Biol. 2015;1273:43–54.
167. Ceroni A, Maass K, Geyer H, et al. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. J Proteome Res. 2008;7:1650–1659.
168. Meitei NS, Apte A, Snovida SI, et al. Automating mass spectrometry-based quantitative glycomics using aminoxy tandem mass tag reagents with SimGlycan. J Proteomics. 2015;127:211–222.
169. Goldberg D, Sutton-Smith M, Paulson J, et al. Automatic annotation of matrix-assisted laser desorption/ionization N-glycan spectra. Proteomics. 2005;5(4):865–875.
170. Yu CY, Mayampurath A, Hu Y, et al. Automated annotation and quantification of glycans using liquid chromatography-mass spectrometry. Bioinformatics. 2013;29(13):1706–1707.
171. Tsai PL, Chen SF. A brief review of bioinformatics tools for glycosylation analysis by mass spectrometry. Mass Spectrom (Tokyo). 2017;6(Spec Iss):S0064.
172. Campbell MP, Nguyen-Khuong T, Hayes CA, et al. Validation of the curation pipeline of UniCarb-DB: building a global glycan reference MS/MS repository. Biochim Biophys Acta. 2014;1844(1 Pt A):108–116.
173. Hayes CA, Karlsson NG, Struwe WB, et al. UniCarb-DB: a database resource for glycomic discovery. Bioinformatics. 2011;27(9):1343–1344.
174. MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010;26(7):966–968.
175. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature Protocols. 2016;11(12):2301–2319.
176. Cooper CA, Gasteiger E, Packer NH. GlycoMod–a software tool for determining glycosylation compositions from mass spectrometric data. Proteomics. 2001;1(2):340–349.
177. Ep G, Rebecchi KR, Dalpathado DS, et al. GlycoPep DB: a tool for glycopeptide analysis using a "Smart Search". Anal Chem. 2007;79(4):1708–1713.
178. Irungu J, Go EP, Dalpathado DS, et al. Simplification of mass spectral analysis of acidic glycopeptides using GlycoPep ID. Anal Chem. 2007;79(8):3065–3074.
179. Hansen JE, Lund O, Tolstrup N, et al. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconj J. 1998;15(2):115–130.
180. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput. 2002;7:310–322.
181. Ozohanics O, Krenyacz J, Ludanyi K, et al. GlycoMiner: a new software tool to elucidate glycopeptide composition. Rapid Commun Mass Spe. 2008;22(20):3245–3254.
182. Bern M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. Curr Protoc Bioinform. 2012;13(20):1–14.
183. Wu Y, Mechref Y, Klouckova I, et al. Mapping site-specific protein N-glycosylations through liquid chromatography/mass spectrometry and targeted tandem mass spectrometry. Rapid Commun Mass Spe. 2010;24(7):965–972.
184. Chalkley RJ, Baker PR. Use of a glycosylation site database to improve glycopeptide identification from complex mixtures. Anal Bioanal Chem. 2017;409(2):571–577.
185. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. Anal Chem. 1997;69(23):4751–4760.
186. Wang J, Qiu S, Chen S, et al. MALDI-TOF MS imaging of metabolites with a N-(1-naphthyl) ethylenediamine dihydrochloride matrix and its application to colorectal cancer liver metastasis. Anal Chem. 2015;87(1):422–430.
187. Uzbekova S, Elis S, Teixeira-Gomes AP, et al. MALDI mass spectrometry imaging of lipids and gene expression reveals differences in fatty acid metabolism between follicular compartments in porcine ovaries. Biology. 2015;4(1):216–236.
188. Touboul D, Brunelle A. MALDI mass spectrometry imaging of lipids and primary metabolites on rat brain sections. Methods Mol Biol. 2015;1203:41–48.
189. Quiason CM, Shahidi-Latham SK. Imaging MALDI MS of dosed brain tissues utilizing an alternative analyte pre-extraction approach. J Am Soc Mass Spectrom. 2015;26(6):967–973.
190. Patel E, Cole LM, Bradshaw R, et al. MALDI-MS imaging for the study of tissue pharmacodynamics and toxicodynamics. Bioanalysis. 2015;7(1):91–101.
191. Gustafsson OJ, Briggs MT, Condina MR, et al. MALDI imaging mass spectrometry of N-linked glycans on formalin-fixed paraffin-embedded murine kidney. Anal Bioanal Chem. 2015;407(8):2127–2139.
192. Barry JA, Groseclose MR, Robichaud G, et al. Assessing drug and metabolite detection in liver tissue by UV-MALDI and IR-MALDESI mass spectrometry imaging coupled to FT-ICR MS. Int J Mass Spectrom. 2015;377:155–448.
193. Zemski Berry KA, Gordon WC, Murphy RC, et al. Spatial organization of lipids in the human retina and optic nerve by MALDI imaging mass spectrometry. J Lipid Res. 2014;55(3):504–515.
194. Rebours V, Le Faouder J, Laouirem S, et al. In situ proteomic analysis by MALDI imaging identifies ubiquitin and thymosin-

beta4 as markers of malignant intraductal pancreatic mucinous neoplasms. Pancreatology. 2014;14(2):117–124.
195. Norris JL, Caprioli RM. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. Chem Rev. 2013;113(4):2309–2342.
196. Aichler M, Walch A. MALDI imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. Lab Invest. 2015;95(4):422–431.
197. Rompp A, Spengler B. Mass spectrometry imaging with high resolution in mass and space. Histochem Cell Biol. 2013;139 (6):759–783.
198. Rompp A, Guenther S, Schober Y, et al. Histology by mass spectrometry: label-free tissue characterization obtained from high-accuracy bioanalytical imaging. Angew Chem Int Ed Engl. 2010;49 (22):3834–3838.
199. Rompp A, Guenther S, Takats Z, et al. Mass spectrometry imaging with high resolution in mass and space (HR(2) MSI) for reliable investigation of drug compound distributions on the cellular level. Anal Bioanal Chem. 2011;401(1):65–73.
200. Anderson DM, Ablonczy Z, Koutalos Y, et al. High resolution MALDI imaging mass spectrometry of retinal tissue lipids. J Am Soc Mass Spectrom. 2014;25(8):1394–1403.
201. Toghi Eshghi S, Yang S, Wang X, et al. Imaging of N-linked glycans from formalin-fixed paraffin-embedded tissue sections using MALDI mass spectrometry. ACS Chem Biol. 2014;9(9):2149–2156.
202. Drake RR, Powers TW, Jones EE, et al. MALDI mass spectrometry imaging of N-linked glycans in cancer tissues. Adv Cancer Res. 2017;134:85–116.
203. Briggs MT, Kuliwaba JS, Muratovic D, et al. MALDI mass spectrometry imaging of N-glycans on tibial cartilage and subchondral bone proteins in knee osteoarthritis. Proteomics. 2016;16(11–12):1736–1741.
204. Powers TW, Neely BA, Shao Y, et al. MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. PloS One. 2014;9(9): e106255.
205. Heijs B, Holst S, Briaire-De Bruijn IH, et al. Multimodal mass spectrometry imaging of N-glycans and proteins from the same tissue section. Anal Chem. 2016;88(15):7745–7753.
206. Everest-Dass AV, Briggs MT, Kaur G, et al. N-glycan MALDI imaging mass spectrometry on formalin-fixed paraffin-embedded tissue enables the delineation of ovarian cancer tissues. Mol Cell Proteomics. 2016;15(9):3003–3016.
207. Holst S, Heijs B, De Haan N, et al. Linkage-specific in situ sialic acid derivatization for N-glycan mass spectrometry imaging of formalin-fixed paraffin-embedded tissues. Anal Chem. 2016;88(11):5904–5913.
208. Pu Y, Ridgeway ME, Glaskin RS, et al. Separation and identification of isomeric glycans by selected accumulation-trapped ion mobility spectrometry-electron activated dissociation tandem mass spectrometry. Anal Chem. 2016;88(7):3440–3443.
209. Liu Y, Clemmer DE. Characterizing oligosaccharides using injected-ion mobility/mass spectrometry. Anal Chem. 1997;69(13):2504–2509.
210. Plasencia MD, Isailovic D, Merenbloom SI, et al. Resolving and assigning N-linked glycan structural isomers from ovalbumin by IMS-MS. J Am Soc Mass Spectrom. 2008;19(11):1706–1715.
211. Vakhrushev SY, Langridge J, Campuzano I, et al. Ion mobility mass spectrometry analysis of human glycourinome. Anal Chem. 2008;80 (7):2506–2513.
212. Hinneburg H, Hofmann J, Struwe WB, et al. Distinguishing N-acetylneuraminic acid linkage isomers on glycopeptides by ion mobility-mass spectrometry. Chem Commun (Camb). 2016;52 (23):4381–4384.
213. Isailovic D, Plasencia MD, Gaye MM, *et al*. Delineating diseases by IMS-MS profiling of serum N-linked glycans. J Proteome Res. 2012;11(2):576–585.
214. De Haan N, Reiding KR, Haberger M, et al. Linkage-specific sialic acid derivatization for MALDI-TOF-MS profiling of IgG glycopeptides. Anal Chem. 2015;87(16):8284–8291.
215. Moh ES, Lin CH, Thaysen-Andersen M, et al. Site-specific N-glycosylation of recombinant pentameric and hexameric human Ig M. J Am Soc Mass Spectrom. 2016;27(7):1143–1155.

1.5 PAPER 2 – HUMAN DISEASE GLYCOMICS: TECHNOLOGY ADVANCES ENABLING PROTEIN GLYCOSYLATION ANALYSIS – PART 2

The immune system has been the subject of frequent and historic study due to it being the cellular response to pathogens, drugs and self. Through the understanding of the biochemical pathways involved in the cellular response and the effects of this cellular response, we can under the mechanism of disease, improve its diagnosis and help develop treatments.

In the following paper, we describe the role of protein-glycosylation in a range of immune system-related diseases (cancer, diabetes, neurological diseases). Although neurological diseases are only recently receiving attention for glycan analysis, glycan structures are increasingly found to be involved in the pathology and mechanisms of disease.

In paper 2, I wrote section 1.2, (Neural Diseases in the Glyco context), prepared Figure 1 and Table 1 and Table 2 and helped edit the manuscript. All figures and tables are compliant with new criteria for displaying glycan structures (Symbol Nomenclature for Glycans[1]). The manuscript design, planning and implementation were discussed by all authors. Overall responsibility for the manuscript was with first author Arun Everest-Dass.

Expert Review of Proteomics

ISSN: 1478-9450 (Print) 1744-8387 (Online) Journal homepage: http://tandfonline.com/loi/ieru20

Human disease glycomics: technology advances enabling protein glycosylation analysis – part 2

Arun V Everest-Dass, Edward S X Moh, Christopher Ashwood, Abdulrahman M M Shathili & Nicolle H Packer

To cite this article: Arun V Everest-Dass, Edward S X Moh, Christopher Ashwood, Abdulrahman M M Shathili & Nicolle H Packer (2018) Human disease glycomics: technology advances enabling protein glycosylation analysis – part 2, Expert Review of Proteomics, 15:4, 341-352, DOI: 10.1080/14789450.2018.1448710

To link to this article: https://doi.org/10.1080/14789450.2018.1448710

Accepted author version posted online: 09 Mar 2018.
Published online: 20 Mar 2018.

EXPERT REVIEW OF PROTEOMICS, 2018
VOL. 15, NO. 4, 341–352
https://doi.org/10.1080/14789450.2018.1448710

Taylor & Francis
Taylor & Francis Group

REVIEW

Human disease glycomics: technology advances enabling protein glycosylation analysis – part 2

Arun V Everest-Dass[a,b,c], Edward S X Moh[a,b], Christopher Ashwood[a,b], Abdulrahman M M Shathili[a,b] and Nicolle H Packer[a,b,c]

[a]Faculty of Science and Engineering, Biomolecular Discovery and Design Research Centre, Macquarie University, Sydney, Australia; [b]ARC Centre for Nanoscale BioPhotonics, Macquarie University, Sydney, Australia; [c]Institute for Glycomics, Griffith University, Gold Coast, Australia

ABSTRACT

Introduction: The changes in glycan structures have been attributed to disease states for several decades. The surface glycosylation pattern is a signature of physiological state of a cell. In this review we provide a link between observed substructural glycan changes and a range of diseases.

Areas covered: We highlight biologically relevant glycan substructure expression in cancer, inflammation, neuronal diseases and diabetes. Furthermore, the alterations in antibody glycosylation in a disease context are described.

Expert commentary: Advances in technologies, as described in Part 1 of this review have now enabled the characterization of specific glycan structural markers of a range of disease states. The requirement of including glycomics in cross-disciplinary *omics* studies, such as genomics, proteomics, epigenomics, transcriptomics and metabolomics towards a systems glycobiology approach to understanding disease mechanisms and management are highlighted.

ARTICLE HISTORY
Received 29 October 2017
Accepted 2 March 2018

KEYWORDS
Glycomics; glycan structure; glycan analysis; mass spectrometry; disease glycomics

1. Association of glycan structural motifs with disease states

Glycosylation of proteins has been known to alter in disease for decades. Given that almost all mammalian cells are covered with glycans, it is no surprise that protein glycosylation plays a role in many human diseases. Changes in cell-surface glycan patterns, often correlated to disease phenotype, can result in specific substructures being displayed at an increased or decreased frequency or even in conveying a change in net charge on the cell surface, both of which could have a significant impact on cell-to-cell interactions. A recent editorial has identified post-translational modifications, including glycosylation, as being useful in understanding dynamic disease mechanisms, when put in context with other omics strategies [1]. The technologies described in Part 1 of the review now enable the determination and relative quantitation of these detailed structural motifs, or glycotopes, that are displayed, usually on the surface of cells, in a multitude of disease states. The obvious differences in expressed glycan structures in diseased cells to those displayed on healthy cells offer opportunities for new biomarkers and therapeutic targets for numerous diseases. Some relevant examples are given here.

1.1. Cancer glycomics

The most studied glycosylation changes over the last decades have been in cancer and have given insight into the significance of glycome changes in most cancers. This has been made possible due to a combination of improved tumor cell biology techniques and advancements in glycomics analysis technologies. Glycan changes observed in cancer have revealed underlying changes in biosynthetic molecular pathways as well as aberrant glycotopes that correlate with more aggressive cancer cell and tumor features, including increased migration, invasion, and metastatic potential, thus providing novel targets for therapeutic intervention [2–7]. In addition to surface glycosylation changes, the secreted glycoproteins from these epithelial cancer cells also carry altered glycans into the blood stream and therefore provide a source of potential new serum biomarkers. The importance of these glycan terminal substructures (Figure 1) from various cancers has been recently reviewed in several publications [5,8–13], and are summarized and updated here.

The most commonly observed aberrant *N*- and *O*-linked protein glycosylation in cancer are changes in the terminal sialylation, fucosylation, and branching. In mucin type *O*-GalNAc glycosylation, the *O*-glycans are mostly observed to be truncated.

Sialylation is predominantly observed at the terminal or nonreducing ends of the glycans of glycoconjugates. The addition of sialic acid is carried out by more than 20 distinct Golgi-resident sialyltransferases (ST) that link sialic acids via their second carbon (C2) to the carbon atom at position C3 (ST3Gal I-VI), C6 (ST6Gal I,II and ST6GalNAc I-VI), or to C8 (ST8Sia I-VI) of a terminal monosaccharide residue, yielding α2,3, α2,6, or α2-8 linkages [14]. An overall increase in sialylation notably of α2–3 and α2,6 linkages on lactosamine

CONTACT Nicolle H Packer nicki.packer@mq.edu.au Department of Chemistry and Biomolecular Sciences, Macquarie University, NSW-2109 North Ryde, Australia

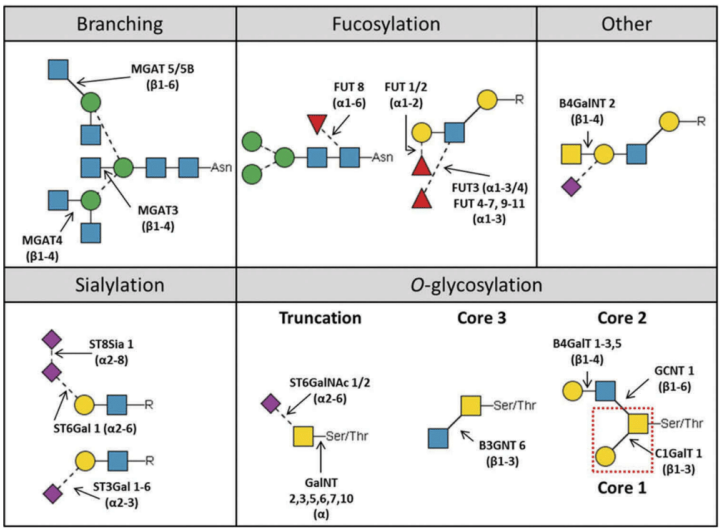


Figure 1. Cell surface protein glycosylation changes, and associated glycosyltransferase genes, observed to alter in cancer. Adapted from Christiansen *et al* [8].

moieties has been associated with several cancers [11]. For example, melanoma cell lines have showed an increased α2,3 sialylation associated with cancer progression of four different melanoma cell lines (WM1552C, WM115, IGR-39, and WM266-4) when evaluated using lectins [15]. Similarly, Sethi *et al*. showed that α2,3 sialylation of *N*-glycans was abundantly identified in the metastatic (LIM1215) and aggressive (LIM2405) colorectal cancer cell lines [16]. In breast cancer, ST3Gal3 and ST3Gal1, the ST associated with the α2,3 sialylation of *N*- and *O*-glycans, respectively, were found to be upregulated [17,18]. On the other hand, Anugraham *et al*. used mass-spectrometry-based analysis to show that α2,6 sialylated *N*-glycans were abundant and unique in both ovarian cancer cell lines and serous ovarian cancer tissue [19,20]. Some other sialylated glycan antigens commonly associated with cancer are the structurally related sialylated Lewis epitopes such as sLewis^a (NeuAc α2–3Galβ1–3(Fucα1–4)GlcNAc) and sLewis^x (NeuAcα2–3Galβ1–4(Fucα1–3)GlcNAc) [21–23]. These terminal glycotopes have been suggested to play a crucial role in cancer metastasis through adhesion of the tumor cells to the endothelium [24–26]. Polysialic acid (PSA) is an α2-8 linked polymer of sialic acid and is a developmentally regulated glycan modification predominantly found in the central nervous system on neuronal cell adhesion molecule. PSA has now been found to be associated with several cancers in modulating cell–cell and cell–matrix adhesion, migration, invasion and metastasis and is strongly associated with poor clinical prognosis [11,27,28].

L-fucose or deoxyhexose is a common constituent of many *N*- and *O*-linked glycans produced by mammalian cells. For example, fucosylation plays important roles in blood group determination, immunological reactions, and signal transduction pathways [29,30]. Altered fucosyltransferases (FUTs) expression have been extensively reported in various cancers [31,32]. In mammals, fucosylation is generally a non-extendable modification and can be subdivided into terminal fucosylation (Lewis blood-group antigens, such as Lewis^x, Lewis^y, Lewis^a and Lewis^b, and the H-antigen) and core fucosylation. The sLewis antigens previously mentioned include the terminal addition of an α1–3 or α1-4 fucose to an α2-3 sialylated type 1 lactosamine or type 2 lactosamine unit, respectively. Increased expression of Lewis^x antigens have been reported in colorectal, breast, and ovarian cancer cell lines [8]. They have generally been associated with higher malignancy and poor prognosis, as previously reported in colon and breast cancer [33–35], but interestingly, Rabassa *et al*. recently reported that Lewis^x expression is associated with a better outcome in patients with head and neck squamous carcinoma [36]. Core fucosylation is the addition of α1-6 fucose residue to the innermost GlcNAc residue of the *N*-glycan chitobiose core by a single fucosyltransferase, FUT8. Core fucosylation of *N*-glycans is significantly upregulated in several cancers including liver, lung, colon and ovarian and breast cancers [10,37–39]. The increase of core fucosylation in certain serum glycoproteins can be used as a diagnostic marker; for example, core fucosylation of α-fetoprotein is an approved biomarker for the early detection of hepatocellular carcinoma (HCC) and is able to distinguish HCC from other liver diseases such as chronic hepatitis and liver cirrhosis [40].

The glycosyltransferase GnT-V catalyzes the addition of a GlcNAc residue to the α1-6 mannose branch of the *N*-glycan core through a β1-6 linkage [41]. Several studies have shown that the increased β1-6 branch and its corresponding MGAT5 gene is involved in cancer growth and metastasis of glioma, colon cancer, and gastric cancer cell lines [42–44]. Also tumors in GnT-V enzyme knockout mice showed decreased growth

and impaired metastasis [45]. This extended branch of β1-6 GlcNAc is further elongated by lactosamine chains to make polylactosamine structures that are high affinity ligands for galectins. Intriguingly, the action of GnT-V is inhibited by the presence of a bisecting GlcNAc [41], the bisecting GlcNAc is transferred to the trimannosyl core of complex or hybrid *N*-glycans on glycoproteins by the β1,4-N-acetylglucosaminyltransferase III (GlcNAcT-III) or MGAT3. Nagae et al. recently used molecular conformational dynamics, crystallography, and NMR analysis to show the major conformational effect that bisecting GlcNAc has on *N*-glycans [46]. They suggested that the addition of this residue restricts the *N*-glycan conformation to a back-fold type, restricting the accessibility of other glycosyltransferases. Generally, bisecting GlcNAc has been suggested to reduce cancer growth and metastatic potential through the regulation of important cell membrane glycoproteins such as EGFR, integrins, and cadherins [47–49]. Other studies have shown that MGAT3 is not always a tumor suppressing factor but can enhance cancer progression [16,50–52]. This is especially the case in high grade ovarian cancer, where an improved overall survival was observed with reduced MGAT3 expression, attributed to epigenetic regulation of this gene[53].

The GalNAc *O*-linked to serine or threonine is the initial sugar added in the *O*-glycan synthesis pathway and is commonly referred to as the Thomsen-nouvelle antigen (Tn). This monosaccharide is usually further extended to a multitude of glycan structures [54]. The sialylation of this antigen (sialyl Tn) effectively truncates any further extension. Tn and sialyl Tn are well-established tumor markers and are mostly correlated with cancer invasion and metastasis [55]. The expression of sialyl Tn is usually low or absent in epithelial cells while is highly abundant in breast, colorectal, pancreatic, ovarian, and endometrial cancers [55,56]. Serum sialyl Tn levels have also been used as a prognostic indicator for epithelial ovarian cancer aggressiveness and metastatic potential [57]. There is a long known strong correlation between sialyl Tn expression and cancer progression but the specific mechanism and effects of sialyl Tn antigen on tumor cells are poorly understood.

1.2. Neuronal diseases in the glyco context

Many diseases that involve neuronal damage have altered glycan substructures, which may act as specific or broad markers for the disease. These glycan structures can be present on the cell surface or presented on proteins secreted by abnormal cells. Proteins secreted into the plasma are often the first place to look for biomarker discovery; however, the central nervous system features cerebrospinal fluid as a unique biological fluid for collection and analysis [58]. This fluid, while important functionally for human health, can also act as a potential source of biomarkers for diseases of the CNS and has been investigated as such [59,60].

In addition to proteins secreted by the brain into the CSF, the brain has its own specialized regions which, evidence suggests, possesses their own unique glycan profile [61]. The brain is a unique organ in terms of its structure and function and this is echoed by the glycans specific for this tissue region. Originally, glycoproteins highly abundant in the brain were studied such as CD24, identifying a range of glycotopes some of which are rarely found elsewhere in the body (Table 1). Common glycan substructures such as Lewis^x, implicated in general inflammation and impaired neurite growth, are accompanied by rarer substructures such as Human Natural Killer 1, a key factor in normal dendritic spine maturation [62].

1.2.1. Impaired differentiation and regeneration of neuronal cells in disease

While these glycotopes are of interest based around brain development and function, more research focus has been oriented toward identifying glycan substructures correlated with disease-related impaired differentiation and regeneration of neuronal cells in the brain (Neuroglycomics). For example, Fogli *et al.*, analyzing the protein glycosylation in the CSF, found four biantennary *N*-glycan diagnostic markers of leukodystrophies correlated with mutations of genes encoding the translation initiation factor, EIF2B [70]. Interestingly, these glycan changes were not echoed in the plasma of the patients, which further emphasizes CSF as an important biological fluid for biomarker discovery.

Impaired differentiation and regeneration mechanisms have ties to developmental disorders such as Amyotrophic Lateral Sclerosis and Huntington's disease. One key glycan of interest in this area is PSA and its role in peripheral nerve regeneration, where a long-term increase in PSA levels in regenerating nerves may favor selective motor target re-innervation [71]. The researchers suggested that controllable polysialyltransferase activity would be an interesting therapeutic approach to the treatment of these diseases.

1.2.2. Huntington's disease

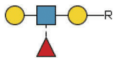
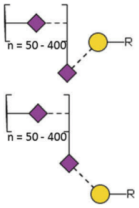
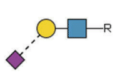
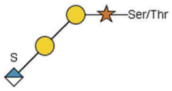
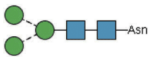
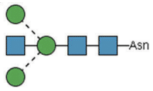
Investigating Huntington's disease, a progressive neurodegenerative disease, Gizaw *et al.* performed, to date, the only comprehensive glycomic profiling of a Huntington's disease mouse model [72]. Both tissue and plasma glycoproteins were analyzed from these mice, identifying an increase in core-fucosylated and bisecting GlcNAc *N*-glycans in the brain tissue of the Huntington's disease model mice compared to control mice. The core-fucosylation motif was found also to be increased in the serum of the Huntington's disease model mice (on biantennary *N*-glycans) as well as having an increase in NeuGc. This study also emphasizes the need for biomarker research in neurodegenerative diseases to move into the field of glycomics.

While there is still a need for further glycomic analysis of the cellular effects of Huntington's disease, Huntingtin, the protein encoded by the Huntington disease gene, has been studied in detail. Although it is an intracellular protein, and does not feature *N*-glycosylation, *O*-GlcNAcylation has been found to play a detrimental role in mutant Huntingtin toxicity in cell and fly models, by preventing the clearance of protein aggregates [73]. The prevention of protein aggregates by *O*-GlcNAc is not disease specific as this effect can also can be found in type II diabetes [74].

1.2.3. Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), much like Huntington's disease, is a neurodegenerative disease that primarily affects motor neurons; however, the exact mechanisms that cause this disease are not as well known. Edri-Brami *et al.* compared the serum of ALS patients with healthy donor controls and found a bi-

Table 1. Glycan substructures commonly observed in diseases of the central nervous system, R: *N*- or *O*-glycan scaffold.

Glycan substructure name	Substructure cartoon	Cell/tissue localization	Implicated in development/disease	Change in abundance
Lewis^X		Neurons [63] CD24 from mouse brain [64] Whole brain [65]	Impaired neurite growth [63] Inflammation [66]	Decrease Increase
Polysialic acid		Microglia [71]	Nerve regeneration [67]	Increase
Alpha 2-3 linked sialic acid		Whole brain [65]	N/A	N/A
Human Natural Killer 1		CD24 from mouse brain [64]	Axon outgrowth, regrowth of motor neurons [62]	Increase
Paucimannosylation		Neurons [68]	Neural stem cell proliferation [68]	Increase
Bisecting GlcNAc		CD24 from mouse brain [64] CSF [69]	Alzheimer's disease [69] Brain developmental disorders [70]	Increased Decreased

antennary glycan of IgG to be upregulated as well as decreased core-fucosylation [75], both of which were found to have functional effects through antibody-dependent cellular cytotoxicity. This demonstrates that profiling of IgG glycosylation is useful for discovering inflammation markers in plasma; however, the specificity of these IgG glycans to ALS has yet to be determined.

1.2.4. Pain

Peripheral nerve regeneration is one of the current chronic pain research areas and the European Union Seventh Framework Programme has identified glycomics as a valuable tool in identifying novel and prognostic biomarkers to predict chronic pain states [76]. As part of this framework programme and in investigating lower back pain, Freiden *et al.* [77] correlated plasma IgG glycan levels with lower back pain and found core-fucosylation to be at significantly lower levels in those suffering lower back pain. The mechanism is thought to be based on core-fucosylation acting as a safety switch to block antibody-dependent cell-mediated cytotoxicity, suggesting core-fucosylation may be an important glycan substructure to monitor for inflammation.

1.2.5. Alzheimer's and Parkinson's diseases

Alzheimer's and Parkinson's disease are both neurodegenerative diseases that typically present in patients later in life. These diseases are therefore hypothesized to be age-related but the cause of Alzheimer's disease, unlike Parkinson's disease, has yet to be established. A review of the proteomic studies on cerebrospinal fluid biomarkers of Alzheimer's disease has recently been published and serves as a useful starting point upon which protein glycosylation can be considered [78]. Similar to low back pain, changes in the *N*-glycosylation of IgG in the plasma of patients with Parkinson's disease featured several glycans as novel biomarkers, with reduced relative abundance of sialylated mono-antennary *N*-glycans [79]. One difficulty in identifying biomarkers or treatment candidates for these diseases is the high correlation with age, where protein glycosylation is known to also alter with

human longevity [80]. *O*-glycosylation analysis has proven most fruitful in the area of Alzheimer's disease, identifying a highly sialylated *O*-glycan structure on amyloid-beta protein which was found to be highly abundant in patients' serum [81,82]. In addition, *N*-glycan profiling identified amyloid-beta to be more glycosylated than in controls, featuring higher relative abundances of bisecting GlcNAc and increased core-fucosylation. Research into treatment methods to follow on from these studies continues to look at enhancing glycosyltransferase expression, particularly MGAT3 which is the enzyme responsible for adding a bisecting GlcNAc, a glycan feature that reportedly is protective in Alzheimer's disease [83,84].

1.3. Glycan features of diabetes

Diabetes is a complex process with several physiological effects in the human body such as increase in glucose levels and insulin resistance; the latter causing a spike in the pro-inflammatory cytokines released [85]. All of these physiological changes have been reported to alter the glycan biosynthetic pathway and the glycome in general [86–90] (Table 2). In the case of inflammation, treating two human pancreatic ductal adenocarcinoma cell lines with several pro-inflammatory cytokines all resulted in higher staining for sialyl Lewis^{x} and sialyl Lewis^{a}. Differences in mRNA expression of ST and FUT, involved in the biosynthesis of these antigens, were determined [90]. TNF-α resulted in an increase in α2-3 ST and α1-2,3 fucosyltransferase mRNA expression which correlated with more sialyl Lewis^{x} and Lewis^{y} lectin staining using immunohistochemistry [90]. Other studies have also reported similar changes in α2-3 ST expression with the proinflammatory cytokine TNF-α [86,91,92]

Previous studies have utilized protein glycosylation as a biomarker for diabetes via detecting the NMR signal of GlcNAc residues on several acute phase glycoproteins such as α-acid glycoprotein, haptoglobin, α1-antitrypsin, α1-antichymotrypsin, and transferrin [94,95]. Since acute phase glycoproteins are also a sign of general inflammation, detailed studies of each protein have been carried out over the years in order to allow for a higher level of accurate diagnosis. For instance, the study of glycans on α-acid glycoprotein showed an increase in α1-3 fucosylation using HPAEC-PAD and immunohistochemistry [96]. This was followed by a study using MALDI-TOF which showed α1-3 fucosylation changes on α-acid glycoprotein in type 2 diabetes [97].

1.3.1. Branched glycan structures

Glucose transporter 2 (GLUT2) is essential for sensing glucose changes in pancreatic B cells and maintaining appropriate insulin secretion. Recently, the elucidation of the molecular mechanism of GLUT2 function has revealed branched β1-4 GlcNAc linkage involvement in binding to Galectin-9 on the B cell surface [98]. A locus for diabetes onset 1q11.5 was seen to encode the gene for GnT-IVa which is the enzyme that catalyzes the transfer of GlcNAc to the core structure of *N*-linked oligosaccharides resulting in multi-antennary structures [98]. A reduction in GnT-IVa expression in pancreatic B cells of islets from humans with type 2 diabetes was evident [99]. Interestingly, knockout GnT-IVa pancreatic B cells showed a significantly shortened half-life time of GLUT2 on the cell surface suggesting a role of branched glycan structures in controlling the stability and residency of the GLUT2 on the B cell surface. Mimicking Galectin-9 binding to GLUT2 resulted in reduction of B cell surface expression levels of GLUT2 [100,101]. Overall, branched glycan structures are important for GLUT2 function and consequently of maintaining appropriate insulin secretion from pancreatic cells.

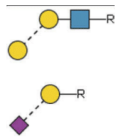
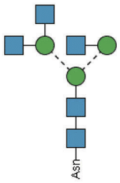
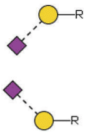
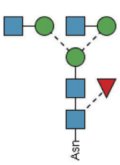
1.3.2. Sialic acid

The human abundant blood plasma chitrotriosidase (CHIT1) protein was extracted from 28 patients with type 2 diabetes and compared with 11 healthy individuals, in order to profile the glycosylation profile by lectin-ELISA [102]. The results showed significant decrease in SNA and MAA lectin binding, suggesting an overall decrease in sialic acid of both α2-6 and α2-3 linkages, respectively. A study of gestational diabetes mellitus has shown a more specific decrease in α2-6 sialic acid linked to the glycans on the pregnancy-related Glycodelin-A. This reduction in α2-6 sialic acid linkage, studied by MALDI-TOF and GC-MS, resulted in defective immune-suppressive activity by Glycodelin-A [103].

1.3.3. Fucosylation

Hepatocyte nuclear factor 1-α (*HNF1α)* has been shown to be a master transcriptional factor of a rare type of diabetes called maturity-onset diabetes of the young (MODY) and there is a strong correlation between fucosyltransferases (*FUT6* and *FUT8)* and *HNF1α* expression in human plasma [82,104]. The relationship between MODY and *HNF1α* has been confirmed by a study that created a fucosylation index (DG9-glycan index) of fucosylated to non-fucosylated tri-antennary structures in plasma to differentiate control and other diabetes types [105]. On the other hand, the hyperglycemic conditions seen in type 2 diabetes result in an overall increase in glucose and UDP-GlcNAc through the hexosamine biosynthetic pathway showing a direct link between protein glycosylation and type 2 diabetes [106,107]. A study of pyridylaminated oligosaccharides using RP-HPLC and MALDI-TOF/TOF of type 2 diabetic mice serum *N*-glycoproteins demonstrated a rise in core-fucosylation that correlated with increased mRNA expression of α1-6 FUT in the liver. The same study and others [97,108] have reported a small, though statistically significant, increase in core fucosylation in serum from human type 2 diabetes patients. However, another study showed an opposing effect in which the serum *N*-glycome of a large cohort (1161 individuals) of healthy and type 2 diabetic subjects showed a significant decrease of a specific core fucosylated biantennary *N*-glycan in type 2 diabetic subjects using DNA-sequencer-aided-fluorophore-assisted carbohydrate electrophoresis [109].

Table 2. The involvement of glycan structures or glycan substructures in diabetes.

Glycan substructure	Observation	Importance	Method of detection	References
	Increase in di-galactose and decrease in α2-3 sialoglycans	Biomarkers for TNF-α induced insulin resistance in adipocytes	Quantitative glycoproteomics for glycosylation site determination of mouse adipocytes with or without TNF-α	[86]
	Increase in α1-3 fucosylation	Biomarker for increase of α-acid glycoprotein, a known acute phase protein	(HPAEC-PAD and immunohistochemistry) as well as (MALDI-TOF and sialidase treatment)	[96,97]
	β1-4 GlcNAc linkage on 1-3 mannose arm of the core structure forming branched structures	Essential for glucose transporter 2 (GLUT2) function. Defect in GLUT2 results in insulin intolerance	ESI-MS/MS of GLUT2. Knockout GnT-IVa in pancreatic B cells and mimicking GLUT2 binding partner (Galectin-9) both decreased GLUT2 life time.	[99,100]
	Decrease in sialic acid specific serum glycoproteins	Biomarker for chitrotiosidase (CHIT-1) and acute phase proteins	Lectin ELISA (SNA and MAA) of CHIT-1 human serum proteins	[102]
	Decrease in α2-6 sialic acid linkage in glycodelin-A	Possible biomarker for gestational diabetes mellitus	Glycodelin-A from amniotic fluid of diabetic pregnant women by lectin-binding, MALDI-TOF, and GC-MS	[103]
	Fucosylation index	Biomarker for maturity-onset diabetes of the young MODY	Lectin-assay	[82]
	Increase of core fucosylation on mice and human serum proteins	Biomarker for diabetic mice using serum proteins	Pyridylaminated oligosaccharides RP-HPLC, MALDI-TOF/TOF on type 2 diabetic mice serum samples and expression of liver α1,6 fucosyltransferase	[93]
	Decrease of di-antennary core fucosylation of human serum proteins	Biomarker for diabetic humans using serum proteins	DNA-sequencer fluorophore-assisted carbohydrate electrophoresis	[109]

Since all of the acute phase glycoproteins are found in blood, studies have been looking for easier obtained body fluids such as saliva, urine, and tears as diagnostic biomarker sources [110,111]. The investigation of basal tears in diabetes and diabetic retinopathy has shown a difference in glycan structures on the tear proteins from diabetic and diabetic

retinopathy patients. Those structures included low abundance *N*- and *O*-glycans including hybrid, complex and the di-sialylated core 2-mucin *O*-glycan [112].

1.4. Antibody glycan structural features

Antibodies play a huge role in all human diseases as a defense mechanism, but can sometimes be the cause of the problem as well, in cases such as autoimmune diseases. The five major types of antibodies in humans (IgG, IgA, IgM, IgE, and IgD, in order of abundance in serum) are all glycosylated proteins. While most of the scientific and commercial interest revolves around the specificity of the antibody–antigen targeting, the relationship between antibodies and their attached glycan counterparts have been shown to be implicated in a range of biological activities such as antibody–receptor binding, inflammation and immune response, and serum half-life [113–117]. The antibody glycans also sometimes act as a biomarker for different biological states such as inflammation (arthritis and vasculitis) and IgA nephropathy [113,115,118–120], but it is not yet known if the changes in the antibody glycan structures are the cause or the effect of the different biological states.

1.4.1. Immunoglobulin G glycosylation

The majority of high throughput analysis is based on IgG, partly due to the high abundance in human serum, ease of purification/isolation, and streamlined analytical methods available for IgG analysis [121–123]. Single monosaccharide differences of the *N*-linked glycans on the single glycosylation site, Asn297, have been reported to alter IgG functions [117,124–126]. Absence of core fucosylation and presence of the bisecting GlcNAc enhances antibody-dependent cellular cytotoxicity (ADCC) and affinity of IgG toward FcγRIII, whereas galactosylation of IgG1 was shown to be involved in anti-inflammatory activity by promoting association of FcγRIIB with the C-type lectin Dectin-1 [126], and enhancing FcγRIIIa binding affinity [127]. Recently, it was also shown that the sialic acid linkage on the IgG glycans could also result in a functional difference; bi-antennary α2-6 linked sialic acid on the glycan on Rituximab resulted in higher ADCC when compared with the α2-3 linked sialic acid form [128]. This study highlights the importance of in-depth glycan structural characterization, information that is more than just knowing the monosaccharide composition of the attached glycan. One recent high-throughput study, with over 5000 patient samples, showed that the relative abundance of several IgG glycan structures significantly changes with age. After a series of statistical modeling and refinements, it was found that the index of three specific glycan structures (core-fucosylated di-galactose, core-fucosylated bisecting, and core-fucosylated bisecting di-galactose glycans) provided a better correlation between chronological and biological age than current biomarkers of age such as telomere length [120].

Another reason that the effects of human IgG glycosylation have been explored more than the other antibody isotypes is the low complexity of the glycosylation on the protein. Most IgG antibodies only contain one *N*-glycosylation site, Asn297, on each heavy chain, which enables direct correlation between released glycan analysis and site-specific glycan analysis. With IgG, technological advances in glycoengineering have enabled the synthesis/production of homogenous, glycan-defined glycoproteins [129–131] for investigation of highly specific glycan mediated structure–function relationships [128,132]. However, the technology to control glycosylation down to the site-specific level on a multiply glycosylated protein is still unavailable. Regulation of site-specific glycosylation has been correlated with protein tertiary and quaternary structure, where certain glycosylation sites are more solvent exposed for better access to glycosylation enzymes [133,134]. For glycosylation sites that are equally accessible, it is not yet known if specific different glycosylation outcomes on a particular site could result in functional differences of the glycoprotein [135,136].

1.4.2. Glycosylation of other antibodies

The other four types of antibodies, IgA, IgM, IgD, and IgE, are all highly glycosylated with multiple *N*- (all) and *O*- (IgA1 and IgD) glycosylation sites [115]. The investigation of site-specific glycosylation changes affecting antibody function has been investigated since the 1990s, predominantly by glycosylation site knockouts (KO) using mutagenesis of the glycosylation sequon [137–145]. From the KO experiments on IgM and IgE, a single glycosylation site was identified to be crucial for effector function; Asn402 of IgM in complement activation and Asn394 of IgE in FcεR binding, both of which are homologous to the Asn297 glycosylation site on IgG. Interestingly, both of these glycosylation sites in IgM and IgE only display high mannose type glycans [145–149]. For IgA1, the consensus on the importance of glycosylation sites has not yet been reached, with evidence suggesting the two *N*-glycosylation site KOs either interfere with [138,140], or do not affect [139] IgA effector functions. These discrepancies could arise from different sources of IgA, different analytical methods, and/or different biological assays. To our knowledge, there has yet to be glycosylation differences relating to IgA2 function. For IgD, not much is known apart from the evidence that deglycosylated IgD impairs the IgD effector function [150], and Asn354 glycosylation is required for IgD secretion [137]. Mucin-type *O*-glycosylation is also present on IgA1 and IgD, located in the hinge region between the Fab and Fc portion of the antibody. Site-specific characterization of *O*-linked glycosylation is a difficult process, due to the lack of consensus sequon, heterogeneity of *O*-glycans, and clustering of multiple *O*-glycan sites amidst a short peptide region [151] so little work has been carried out because of this analytical difficulty that has yet to be resolved.

2. Expert commentary

Improved analytical technologies and innovative approaches have provided to a large extent the unraveling of the complex glycan code. This has enabled the identification of specific glycan substructures directly involved in disease manifestations or their progression in several areas including oncology,

neurobiology, diabetes, and immunology. Glycosylation changes observed in diseases are becoming recognized as increasingly valuable in understanding disease development, progression, and treatment.

Although genomic interpretations of the glycosylation machinery can provide a framework encompassing a cell's glycosylation profile, it lacks information about the complex interplay between transferases, glycosidases, nucleotide sugar donors, and epigenetic regulators that are involved in the generation of highly ordered, regulated, and conserved diverse glycan substructures observed in physiology and disease. Over the last decade many glycan-mediated disease associations have been suggested, some of which have been covered in this review. There is increased interest in new therapeutic interventions involving one or more of the many components of the glycosylation machinery, including the glycoengineering of recombinant antibodies. To fully understand the role of aberrant glycosylation changes observed in diseases, a systems-biology approach is required. High-throughput technologies that facilitate the analysis of complex glycan structures (as described in Part 1 of this review), detection methods to identify their interacting partners and innovative hyphenation of existing techniques are required to address these questions.

The simultaneous characterization of both the glycan (substructure) moieties and their protein carriers is also an added advantage especially in the discovery of biomarkers for diseases. This brings an added dimension and can increase the sensitivity associated with several disease biomarkers where both a protein and associated glycosylation change can be used to reduce the false positives. Similarly, a systems-glycobiology approach including the genomics, proteomics, epigenomics, transcriptomics, and metabolomics, all of which are involved in the cellular glycosylation machinery, will benefit the understanding and monitoring of disease mechanisms.

3. Five-year view: a speculative viewpoint on how the field will evolve in 5 years' time

Although several of the current disease biomarkers are glycoproteins (Table 3), diagnosis is often based on concentration of the glycoprotein. It is now clearly evident that incorporating the observed glycan changes in diseases with their protein carriers would dramatically improve the specificity and sensitivity of the current and future glycoprotein markers. With current advancements in analytical technology, particularly in liquid chromatography and mass spectrometry, glycomics can measure the macro- and micro-heterogeneity of these disease glycoprotein biomarkers both qualitatively and quantitatively from complex samples such as serum. The combination of measuring both protein and glycan alterations to glycoproteins in body fluids as diagnostics, and the engineering of therapeutic immunoglobulins to mimic the glycosylation changes that are observed in naturally occurring ADCC responses, in particular present attractive targets in this regard. The continued development and simplification of reliable, robust, and sensitive analysis workflows is essential to transform these technologies into a clinical setting.

The unique glycan substructure changes observed in diseases also provide interesting therapeutic targets for monoclonal antibody (mAb) production. Anti-glycan mAbs have a multifunctional role, as they can be used to detect/locate tumor distribution as well as to inhibit cell proliferation and cause cell death. In addition, although currently it is quite challenging to produce potent mAbs targeting tumor glycans, anti-glycan mAbs could be used as carrier molecules for nanoparticles loaded with drugs or imaging probes.

Table 3. List of current clinically relevant glycoprotein biomarkers.

Glycoprotein	Marker	Type of disease	Clinical application
α-Fetoprotein	Protein concentration and core fucosylation	Hepatoma, testicular cancers	Diagnosis, staging, detection, recurrence, and monitoring
Cancer antigen 125 (CA125, MUC16)	Protein concentration	Ovarian cancer	Detection, monitoring, and recurrence
Human epididymis protein 4 (HE 4)	Protein concentration	Ovarian cancer	Detection, monitoring, and recurrence
Cancer antigen 15-3 (MUC 1)	Sialyl Tn	Breast cancer	Monitoring
Carbohydrate antigen 19-9 (CA19-9)	Sialyl Lewis[a]	Pancreatic, ovarian cancers	Monitoring
Carcinoembryonic antigen (CEA)	Protein concentration	Gastric, colon, pancreatic, breast, and lung cancers	Detection, monitoring, and recurrence
Human epidermal growth factor receptor 2 (HER 2)	Protein concentration	Breast cancer	Choice of therapy
Human chorionic gonadotropin (hCG)	Protein concentration	Testicular, ovarian cancers	Diagnosis, staging, detection, recurrence, and monitoring
Fibrinogen degradation product (FDP)	Protein concentration	Bladder cancer	Detection and monitoring
Thyroglobulin	Protein concentration	Thyroid cancer	Monitoring and recurrence
Prostate-specific antigen (PSA)	Protein concentration	Prostate cancer	Detection, monitoring, and recurrence
Carbohydrate antigen 72–4 (CA72-4)	Sialyl Tn	Gastric cancer	Diagnosis, staging, detection, recurrence, and monitoring
Acute phase glycoproteins (α-acid glycoprotein, haptogloblin, α1-antitrypsin, α1-antichymotrypsin and transferrin)	Protein concentration	Inflammation such as in the case of diabetes	Detection and monitoring
Glucose transporter 2 (GLUT2)	Branching (Lower GLUT2 half-life resulted from GnT-IVa reduction in diabetes)	Type 2 diabetes	Detection and monitoring
Glycodelin-A	α2-6 sialic acid	Gestational diabetes mellitus	Detection and monitoring

This review does not aim to provide an exhaustive analysis of all glycan substructure-mediated changes in diseases, but rather to describe some crucial examples that demonstrate the possible impact of targeting protein glycosylation changes in disease. A major challenge from the beginning in glycomics has been the robust, reproducible, analytical task of deciphering the macro- and micro-heterogeneity changes at the glycoproteome level in an endogenous biological context. With the described (in Part 1) extraordinary analytical advances over the last decade, this now seems a possibility.

Key issues

- Changes in cell-surface glycan patterns in disease result in specific glycan substructures being displayed at an increased or decreased frequency on the cell surface, which could have a significant impact on cell-to-cell interactions and communications.
- The differences in glycan structures observed in diseased cells to those displayed on healthy cells offer opportunities for new biomarkers and therapeutic targets for numerous diseases.
- Glycan changes observed in cancer reveal aberrant glycotopes that correlate with more aggressive cancer cells and tumor features, including increased migration, invasion, and metastatic potential, thus providing novel targets for therapeutic intervention.
- Diseases that involve neuronal damage display unique glycotopes that may act as specific or broad markers for neuronal diseases.
- All of the physiological changes associated with diabetes have been reported to alter the glycan biosynthetic pathway and the glycome.
- Antibody glycosylation can differentiate between different biological states, such as potently altering the inflammatory profile and effector functions of the disease related antibodies.
- Combining glycomics with genomics, proteomics, epigenomics, transcriptomics and metabolomics will provide a systems-glycobiology approach to understanding many diseases.

Declartion of interest

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed. Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Funding

E.S.X Moh and A.M.M Shathili are funded by the International Macquarie University Research Excellence Scholarship. C. Ashwood is funded by the Australian Postgraduate Award. This work was supported by funding from the Australian Research Council Centre for Nanoscale BioPhotonics.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Zahedi RP. Joining forces: studying multiple post-translational modifications to understand dynamic disease mechanisms. Expert Rev Proteomics. 2016;13(12):1055–1057.
2. Kailemia MJ, Park D, Lebrilla CB. Glycans and glycoproteins as specific biomarkers for cancer. Anal Bioanal Chem. 2017;409(2):395–410.
3. Vankemmelbeke M, Chua JX, Durrant LG. Cancer cell associated glycans as targets for immunotherapy. Oncoimmunology. 2016;5(1):e1061177.
4. Nardy AF, Freire-de-Lima L, Freire-de-Lima CG, et al. The sweet side of immune evasion: role of glycans in the mechanisms of cancer progression. Front Oncol. 2016;6(54).
•• **This review discusses the role of glycans in fundamental mechanisms controlling cancer development and progression, and their potential applications in oncology.**
5. Taniguchi N, Kizuka Y. Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. Adv Cancer Res. 2015;126(11–51).
6. Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. Biochim Biophys Acta. 2012;1820(9):1347–1353.
7. Lau KS, Dennis JW. N-glycans in cancer progression. Glycobiology. 2008;18(10):750–760.
8. Christiansen MN, Chik J, Lee L, et al. Cell surface protein glycosylation in cancer. Proteomics. 2014;14(4–5):525–546.
•• **Comprehensive review of glycan changes observed in cancer.**
9. Almeida A, Kolarich D. The promise of protein glycosylation for personalised medicine. Biochim Biophys Acta. 2016;1860(8):1583–1595.
10. Zhao YY, Takahashi M, Gu JG, et al. Functional roles of N-glycans in cell signaling and cell adhesion in cancer. Cancer Sci. 2008;99(7):1304–1310.
11. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 2015;15(9):540–555.
12. Oliveira-Ferrer L, Legler K, Milde-Langosch K. Role of protein glycosylation in cancer metastasis. Semin Cancer Biol. 2017;44:141–152.
13. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. Annu Rev Pathol. 2015;10(473–510).
14. Dall'Olio F, Malagolini N, Trinchera M, et al. Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression. Biochim Biophys Acta. 2014;1840(9):2752–2764.
15. Kolasinska E, Przybylo M, Janik M, et al. Towards understanding the role of sialylation in melanoma progression. Acta Biochim Pol. 2016;63(3):533–541.
16. Sethi MK, Thaysen-Andersen M, Smith JT, et al. Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and alpha-2,3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. J Proteome Res. 2014;13(1):277–288.
17. Recchi MA, Hebbar M, Hornez L, et al. Multiplex reverse transcription polymerase chain reaction assessment of sialyltransferase expression in human breast cancer. Cancer Res. 1998;58(18):4066–4070.
18. Hebbar M, Krzewinski-Recchi MA, Hornez L, et al. Prognostic value of tumoral sialyltransferase expression and circulating E-selectin concentrations in node-negative breast cancer patients. Int J Biol Markers. 2003;18(2):116–122.
19. Anugraham M, Jacob F, Nixdorf S, et al. Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: glycan structures reflect gene expression and DNA methylation status. Molecul Cell Prot. 2014;13(9):2213–2232.
20. Anugraham M, Jacob F, Everest-Dass AV, et al. Tissue glycomics distinguish tumour sites in women with advanced serous adenocarcinoma. Mol Oncol. 2017;11(11):1595–1615.
21. Trinchera M, Aronica A, Dall'Olio F. Selectin ligands sialyl-Lewis a and sialyl-Lewis X in gastrointestinal cancers. Biology. 2017;6(1).
22. Liang JX, Liang Y, Gao W. Clinicopathological and prognostic significance of sialyl Lewis X overexpression in patients with cancer: a meta-analysis. Onco Targets Ther. 2016;9:3113–3125.
23. Borzym-Kluczyk M, Radziejewska I. Changes of the expression of Lewis blood group antigens in glycoproteins of renal cancer tissues. Acta Biochim Pol. 2013;60(2):223–226.

24. Ugorski M, Laskowska A. Sialyl Lewis(a): a tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells. Acta Biochim Pol. 2002;49(2):303–311.
25. Kannagi R. Carbohydrate antigen sialyl Lewis a–its pathophysiological significance and induction mechanism in cancer progression. Chang Gung Med J. 2007;30(3):189–209.
26. Jeschke U, Mylonas I, Shabani N, et al. Expression of sialyl Lewis X, sialyl Lewis A, E-cadherin and cathepsin-D in human breast cancer: immunohistochemical analysis in mammary carcinoma in situ, invasive carcinomas and their lymph node metastasis. Anticancer Res. 2005;25(3A):1615–1622.
27. Wang X, Li X, Zeng YN, et al. Enhanced expression of polysialic acid correlates with malignant phenotype in breast cancer cell lines and clinical tissue samples. Int J Mol Med. 2016;37(1):197–206.
28. Elkashef SM, Allison SJ, Sadiq M, et al. Polysialic acid sustains cancer cell survival and migratory capacity in a hypoxic environment. Sci Rep. 2016;6(33026).
29. Miyoshi E, Moriwaki K, Nakagawa T. Biological function of fucosylation in cancer biology. J Biochem. 2008;143(6):725–729.
30. Schneider M, Al-Shareffi E, Haltiwanger RS. Biological functions of fucose in mammals. Glycobiology. 2017;27(7):601–618.
31. Miyoshi E, Moriwaki K, Terao N, et al. Fucosylation is a promising target for cancer diagnosis and therapy. Biomolecules. 2012;2(1):34–45.
32. Moriwaki K, Miyoshi E. Fucosylation and gastrointestinal cancer. World J Hepatol. 2010;2(4):151–161.
33. Nakagoe T, Fukushima K, Nanashima A, et al. Expression of Lewis (a), sialyl Lewis(a), Lewis(x) and sialyl Lewis(x) antigens as prognostic factors in patients with colorectal cancer. Can J Gastroenterol. 2000;14(9):753–760.
34. Nakagoe T, Fukushima K, Itoyanagi N, et al. Expression of ABH/Lewis-related antigens as prognostic factors in patients with breast cancer. J Cancer Res Clin Oncol. 2002;128(5):257–264.
35. Koh YW, Lee HJ, Ahn JH, et al. Expression of Lewis X is associated with poor prognosis in triple-negative breast cancer. Am J Clin Pathol. 2013;139(6):746–753.
36. Rabassa ME, Pereyra A, Pereyra L, et al. Lewis x antigen is associated to head and neck squamous cell carcinoma survival. Pathol Oncol Res. 2017.
37. Liu YC, Yen HY, Chen CY, et al. Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. Proc Natl Acad Sci U S A. 2011;108(28):11332–11337.
38. Potapenko IO, Haakensen VD, Luders T, et al. Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. Mol Oncol. 2010;4(2):98–118.
39. Abbott KL, Nairn AV, Hall EM, et al. Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. Proteomics. 2008;8(16):3210–3220.
40. Sato Y, Nakata K, Kato Y, et al. Early recognition of hepatocellular carcinoma based on altered profiles of alpha-fetoprotein. N Engl J Med. 1993;328(25):1802–1806.
41. Kizuka Y, Taniguchi N. Enzymes for N-glycan branching and their genetic and nongenetic regulation in cancer. Biomolecules. 2016;6(2).
42. Yamamoto H, Swoger J, Greene S, et al. Beta1,6-N-acetylglucosamine-bearing N-glycans in human gliomas: implications for a role in regulating invasivity. Cancer Res. 2000;60(1):134–142.
43. Kim YS, Ahn YH, Song KJ, et al. Overexpression and beta-1,6-N-acetylglucosaminylation-initiated aberrant glycosylation of TIMP-1: a "double whammy" strategy in colon cancer progression. J Biol Chem. 2012;287(39):32467–32478.
44. Wang X, He H, Zhang H, et al. Clinical and prognostic implications of beta1, 6-N-acetylglucosaminyltransferase V in patients with gastric cancer. Cancer Sci. 2013;104(2):185–193.
45. Granovsky M, Fata J, Pawling J, et al. Suppression of tumor growth and metastasis in Mgat5-deficient mice. Nat Med. 2000;6(3):306–312.
46. Nagae M, Kanagawa M, Morita-Matsumoto K, et al. Atomic visualization of a flipped-back conformation of bisected glycans bound to specific lectins. Sci Rep. 2016;6(22973).
47. Hoja-Lukowicz D, Przybylo M, Duda M, et al. On the trail of the glycan codes stored in cancer-related cell adhesion proteins. Biochim Biophys Acta. 2017;1861(1 Pt A):3237–3257.
48. Miwa HE, Song Y, Alvarez R, et al. The bisecting GlcNAc in cell growth control and tumor progression. Glycoconj J. 2012;29(8–9):609–618.
49. Song Y, Aglipay JA, Bernstein JD, et al. The bisecting GlcNAc on N-glycans inhibits growth factor signaling and retards mammary tumor progression. Cancer Res. 2010;70(8):3361–3371.
50. Yang X, Tang J, Rogler CE, et al. Reduced hepatocyte proliferation is the basis of retarded liver tumor progression and liver regeneration in mice lacking N-acetylglucosaminyltransferase III. Cancer Res. 2003;63(22):7753–7759.
51. Yang X, Bhaumik M, Bhattacharyya R, et al. New evidence for an extra-hepatic role of N-acetylglucosaminyltransferase III in the progression of diethylnitrosamine-induced liver tumors in mice. Cancer Res. 2000;60(12):3313–3319.
52. Allam H, Johnson BP, Zhang M, et al. The glycosyltransferase GnT-III activates Notch signaling and drives stem cell expansion to promote the growth and invasion of ovarian cancer. J Biol Chem. 2017;292(39):16351–16359.
53. Kohler RS, Anugraham M, Lopez MN, et al. Epigenetic activation of MGAT3 and corresponding bisecting GlcNAc shortens the survival of cancer patients. Oncotarget. 2016;7(32):51674–51686.
54. Springer GF. T and Tn, general carcinoma autoantigens. Science. 1984;224(4654):1198–1206.
55. Schultz MJ, Swindall AF, Bellis SL. Regulation of the metastatic cell phenotype by sialylated glycans. Cancer Metastasis Rev. 2012;31(3–4):501–518.
56. Pearce OM, Laubli H. Sialic acids in cancer biology and immunity. Glycobiology. 2016;26(2):111–128.
57. Kobayashi H, Terao T, Kawashima Y. Clinical evaluation of circulating serum sialyl Tn antigen levels in patients with epithelial ovarian cancer. J Clin Oncol. 1991;9(6):983–987.
58. Kroksveen AC, Opsahl JA, Aye TT, et al. Proteomics of human cerebrospinal fluid: discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics. J Proteomics. 2011;74(4):371–388.
59. Garden GA, Campbell BM. Glial biomarkers in human central nervous system disease. Glia. 2016;64(10):1755–1771.
60. Hoshi K, Matsumoto Y, Ito H, et al. A unique glycan-isoform of transferrin in cerebrospinal fluid: a potential diagnostic marker for neurological diseases. Biochim Biophys Acta. 2017;1861(10):2473–2478.
61. Ji IJ, Hua S, Shin DH, et al. Spatially-resolved exploration of the mouse brain glycome by tissue glyco-capture (TGC) and nano-LC/MS. Anal Chem. 2015;87(5):2869–2877.
62. Kizuka Y, Oka S. Regulated expression and neural functions of human natural killer-1 (HNK-1) carbohydrate. Cell Mol Life Sci. 2012;69(24):4135–4147.
63. Gouveia R, Schaffer L, Papp S, et al. Expression of glycogenes in differentiating human NT2N neurons. Downregulation of fucosyltransferase 9 leads to decreased Lewisx levels and impaired neurite outgrowth. Biochimica Biophys Acta Genl Sub. 2012;1820(12):2007–2019.
64. Bleckmann C, Geyer H, Reinhold V, et al. Glycomic analysis of N-linked carbohydrate epitopes from CD24 of mouse brain. J Proteome Res. 2009;8(2):567–582.
65. Hoffmann A, Nimtz M, Wurster U, et al. Carbohydrate structures of ß-trace protein from human cerebrospinal fluid: evidence for "brain-type" N-glycosylation. J Neurochem. 1994;63:2185–2196.
66. Telford JE, Doherty MA, Tharmalingam T, et al. Discovering new clinical markers in the field of glycomics. Biochem Soc Trans. 2011;39(1):327–330.
67. Schnaar RL, Gerardy-Schahn R, Hildebrandt H. Sialic acids in the brain: gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. Physiol Rev. 2014;94(461–518).
68. Dahmen AC, Fergen MT, Laurini C, et al. Paucimannosidic glycoepitopes are functionally involved in proliferation of neural progenitor cells in the subventricular zone. Glycobiology. 2015;25(8):869–880.

69. Palmigiano A, Barone R, Sturiale L, et al. CSF N-glycoproteomics for early diagnosis in Alzheimer's disease. J Proteomics. 2016;131:29–37.
70. Fogli A, Merle C, Roussel V, et al. CSF N-glycan profiles to investigate biomarkers in brain developmental disorders: application to leukodystrophies related to eIF2B mutations. PloS One. 2012;7(8):1–10.
71. Werneburg S, Mühlenhoff M, Stangel M, et al. Polysialic acid on SynCAM 1 in NG2 cells and on neuropilin-2 in microglia is confined to intracellular pools that are rapidly depleted upon stimulation. Glia. 2015;63(7):1240–1255.
72. Gizaw ST, Koda T, Amano M, et al. A comprehensive glycome profiling of Huntington's disease transgenic mice. Biochimica Biophys Acta Genl Sub. 2015;1850(9):1704–1718.
73. Portelius E, Brinkmalm G, Pannee J, et al. Proteomic studies of cerebrospinal fluid biomarkers of Alzheimer's disease: an update. Expert Rev Proteomics. 2017;14(11):1007–1020.
74. Kumar A, Singh PK, Parihar R, et al. Decreased O-linked GlcNAcylation protects from cytotoxicity mediated by huntingtin exon1 protein fragment. J Biol Chem. 2014;289(19):13543–13553.
75. Edri-Brami M, Rosental B, Hayoun D, et al. Glycans in sera of amyotrophic lateral sclerosis patients and their role in killing neuronal cells. PloS One. 2012;7(5):1–15.
76. Kringel D, Lötsch J. Pain research funding by the European Union Seventh Framework Programme. Eur J Pain. 2015;19(5):595–600.
77. Freidin MB, Keser T, Gudelj I, et al. The association between low back pain and composition of IgG glycome. Sci Rep. 2016 May 27;6:26815. doi:10.1038/srep26815. PMID:27229623.
78. Marsh SA, Powell PC, Dell'italia LJ, et al. Cardiac O-GlcNAcylation blunts autophagic signaling in the diabetic heart. Life Sci. 2013;92(11):648–656.
79. Russell AC, Šimurina M, Garcia MT, et al. The N-glycosylation of immunoglobulin G as a novel biomarker of Parkinson's disease. Glycobiology. 2017;27(5):501–510.
80. Ruhaak LR, Uh HW, Beekman M, et al. Decreased levels of bisecting GLcNAc glycoforms of IgG are associated with human longevity. PloS One. 2010;5(9):1–8.
81. Chen CC, Engelborghs S, Dewaele S, et al. Altered serum glycomics in Alzheimer disease: a potential blood biomarker? Rejuvenation Res. 2010;13(4):439–444.
82. Majidi S, Fouts A, Pyle L, et al. Can Biomarkers Help Target Maturity-Onset Diabetes of the Young Genetic Testing in Antibody-Negative Diabetes? Diabetes Technol Ther. 2018 Feb;20(2):106–112. doi:10.1089/dia.2017.0317. PMID:29355436.
83. Fiala M, Mahanian M, Rosenthal M, et al. MGAT3 mRNA: a biomarker for prognosis and therapy of Alzheimer's disease by vitamin D and curcuminoids. J Alzheimers Dis. 2011;25(1):135–144.
84. Fiala M, Porter V. Delineating Alzheimer's disease progression with MGAT3, a biomarker for improved prognosis and personalized therapy. Biomark Med. 2011;5(5):645–647.
85. Hotamisligil G, Shargill N, Spiegelman B. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. Science. 1993;259(5091):87–91.
86. Parker BL, Thaysen-Andersen M, Fazakerley DJ, et al. Terminal galactosylation and sialylation switching on membrane glycoproteins upon TNF-alpha-induced insulin resistance in adipocytes. Mol Cell Proteomics. 2016;15(1):141–153.
87. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. Cell. 2006;126(5):855–867.
88. Rellier N, Ruggiero D, Lecomte M, et al. Advanced glycation end products induce specific glycoprotein alterations in retinal microvascular cells. Biochem Biophys Res Commun. 1997;235(2):281–285.
89. Rellier N, Ruggiero-Lopez D, Lecomte M, et al. In vitro and in vivo alterations of enzymatic glycosylation in diabetes. Life Sci. 1999;64(17):1571–1583.
90. Bassagañas S, Allende H, Cobler L, et al. Inflammatory cytokines regulate the expression of glycosyltransferases involved in the biosynthesis of tumor-associated sialylated glycans in pancreatic cancer cell lines. Cytokine. 2015;75(1):197–206.
91. Majuri ML, Pinola M, Niemelä R, et al. α2, 3-Sialyl and α1, 3-fucosyltransferase-dependent synthesis of sialyl Lewis x, an essential oligosaccharide present on L-selectin counterreceptors, in cultured endothelial cells. Eur J Immunol. 1994;24(12):3205–3210.
92. Majuri ML, Niemelä R, Tiisala S, et al. Expression and function of α2, 3-sialyl-and α, 3/l, 4-fucosyltransferases in colon adenocarcinoma cell lines: role in synthesis of E-selectin counter-receptors. Int J Cancer. 1995;63(4):551–559.
93. Itoh N, Sakaue S, Nakagawa H, et al. Analysis of N-glycan in serum glycoproteins from db/db mice and humans with type 2 diabetes. Am J Physiol Endocrinol Metab. 2007;293(4):E1069–1077.
94. Schmidt MI, Duncan BB, Sharrett AR, et al. Markers of inflammation and prediction of diabetes mellitus in adults (atherosclerosis risk in communities study): a cohort study. The Lancet. 1999;353(9165):1649–1652.
95. Bell JD, Brown JC, Nicholson JK, et al. Assignment of resonances for 'acute-phase'glycoproteins in high resolution proton NMR spectra of human blood plasma. FEBS Letters. 1987;215(2):311–315.
96. Poland DCW, Schalkwijk CG, Stehouwer CDA, et al. Increased α3-fucosylation of α1-acid glycoprotein in type I diabetic patients is related to vascular function. Glycoconj J. 2001;18(3):261–268.
97. Higai K, Azuma Y, Aoki Y, et al. Altered glycosylation of α1-acid glycoprotein in patients with inflammation and diabetes mellitus. Clinica Chimica Acta. 2003;329(1–2):117–125.
98. Ohtsubo K. Glyco-predisposing factor of diabetes. In: Suzuki T, Ohtsubo K, Taniguchi N, editors. Sugar chains: decoding the functions of glycans. TokJapan, yo: Springer; 2015. p. 209–218.
99. Ohtsubo K, Chen MZ, Olefsky JM, et al. Pathway to diabetes through attenuation of pancreatic beta cell glycosylation and glucose transport. Nat Med. 2011;17(9):1067–1075.
100. Ohtsubo K, Takamatsu S, Minowa MT, et al. Dietary and genetic control of glucose transporter 2 glycosylation promotes insulin secretion in suppressing diabetes. Cell. 2005;123(7):1307–1321.
101. Ohtsubo K, Takamatsu S, Gao C, et al. N-glycosylation modulates the membrane sub-domain distribution and activity of glucose transporter 2 in pancreatic beta cells. Biochem Biophys Res Commun. 2013;434(2):346–351.
102. Żurawska-Płaksej E, Kratz EM, Ferens-Sieczkowska M, et al. Changes in glycosylation of human blood plasma chitotriosidase in patients with type 2 diabetes. Glycoconj J. 2016;33(1):29–39.
103. Lee C-L, Chiu PCN, Pang P-C, et al. Glycosylation failure extends to glycoproteins in gestational diabetes mellitus: evidence from reduced α2-6 sialylation and impaired immunomodulatory activities of pregnancy-related glycodelin-A. Diabetes. 2011;60(3):909–917.
104. Yamagata K, Oda N, Kaisaki PJ, et al. Mutations in the hepatocyte nuclear factor-1[alpha] gene in maturity-onset diabetes of the young (MODY3). Nature. 1996;384(6608):455–458.
105. Thanabalasingham G, Huffman JE, Kattla JJ, et al. Mutations in HNF1A result in marked alterations of plasma glycan profile. Diabetes. 2013;62(4):1329–1337.
106. Marshall S, Bacote V, Traxinger R. Discovery of a metabolic pathway mediating glucose-induced desensitization of the glucose transport system. Role of hexosamine biosynthesis in the induction of insulin resistance. J Biol Chem. 1991;266(8):4706–4712.
107. Buse MG. Hexosamines, insulin resistance, and the complications of diabetes: current status. Am J Physiol Endocrinol And Metabolism. 2006;290(1):E1–E8.
108. McMillan DE. Elevation of glycoprotein fucose in diabetes mellitus. Diabetes. 1972;21(8):863–871.
109. Testa R, Vanhooren V, Bonfigli AR, et al. N-Glycomic changes in serum proteins in type 2 diabetes mellitus correlate with complications and with metabolic syndrome parameters. PloS One. 2015;10(3):e0119983.
110. Rao PV, Laurie A, Bean ES, et al. Salivary protein glycosylation as a noninvasive biomarker for assessment of glycemia. J Diabetes Sci Technol. 2015;9(1):97–104.
111. Brownlee M, Vlassara H, Cerami A. Measurement of glycosylated amino acids and peptides from urine of diabetic patients using affinity chromatography. Diabetes. 1980;29(12):1044–1047.

112. Nguyen-Khuong T, Everest-Dass AV, Kautto L, et al. Glycomic characterization of basal tears and changes with diabetes and diabetic retinopathy. Glycobiology. 2015;25(3):269–283.
113. Epp A, Sullivan KC, Herr AB, et al. Immunoglobulin glycosylation effects in allergy and immunity. Curr Allergy Asthma Rep. 2016;16(11):79.
114. K-T S, Anthony R. Antibody glycosylation and inflammation. Antibodies. 2013;2(3):392.
• **This review discusses the contribution of the Fc glycan to IgG antibody effector functions and *in vivo* regulation of the antibody glycosylation.**
115. Maverakis E, Kim K, Shimoda M, et al. Glycans in the immune system and the altered glycan theory of autoimmunity: a critical review. J Autoimmun. 2015;57:1–13.
116. Jefferis R. Glycosylation as a strategy to improve antibody-based therapeutics. Nat Rev Drug Discov. 2009;8(3):226–234.
117. Hudak JE, Bertozzi CR. Glycotherapy: new advances inspire a reemergence of glycans in medicine. Chem Biol. 2014;21(1):16–37.
118. Takahashi K, Raska M, Stuchlova Horynova M, et al. Enzymatic sialylation of IgA1 O-glycans: implications for studies of IgA nephropathy. PloS One. 2014;9(2):e99026.
119. Gornik O, Lauc G. Glycosylation of serum proteins in inflammatory diseases. Dis Markers. 2008;25(4–5):267–278.
120. Krištić J, Vučković F, Menni C, et al. Glycans are a novel biomarker of chronological and biological ages. Journals Gerontol Ser A: Biol Sci Med Sci. 2014;69(7):779–789.
121. Dotz V, Haselberg R, Shubhakar A, et al. Mass spectrometry for glycosylation analysis of biopharmaceuticals. In: TrAC trends in analytical chemistry. 2015;1-9. Available from: https://www.sciencedirect.com/science/article/pii/S0165993615001995
122. Jansen BC, Reiding KR, Bondt A, et al. MassyTools: a high-throughput targeted data processing tool for relative quantitation and quality control developed for glycomic and glycoproteomic MALDI-MS. J Proteome Res. 2015;14(12):5088–5098.
123. Pucic M, Knezevic A, Vidic J, et al. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. Molecul Cell Prot. 2011;10(10):M111.010090.
124. Ratner M. Genentech's glyco-engineered antibody to succeed Rituxan. Nat Biotechnol. 2014;32(1):6–7.
125. Ferrara C, Grau S, Jager C, et al. Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. Proc Natl Acad Sci U S A. 2011;108(31):12669–12674.
126. Karsten CM, Pandey MK, Figge J, et al. Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of Fc[gamma]RIIB and dectin-1. Nat Med. 2012;18(9):1401–1406.
127. Kurogochi M, Mori M, Osumi K, et al. Glycoengineered monoclonal antibodies with homogeneous glycan (M3, G0, G2, and A2) using a chemoenzymatic approach have different affinities for fcγriiia and variable antibody-dependent cellular cytotoxicity activities. PloS One. 2015;10(7):e0132848.
128. Lin CW, Tsai MH, Li ST, et al. A common glycan structure on immunoglobulin G for enhancement of effector functions. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(34):10611–10616.
129. Goodfellow JJ, Baruah K, Yamamoto K, et al. An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. J Am Chem Soc. 2012;134(19):8030–8033.
130. Zou G, Ochiai H, Huang W, et al. Chemoenzymatic synthesis and Fcgamma receptor binding of homogeneous glycoforms of antibody Fc domain. Presence of a bisecting sugar moiety enhances the affinity of Fc to FcgammaIIIa receptor. J Am Chem Soc. 2011;133(46):18975–18991.
131. Zha D. Glycoengineered Pichia-based expression of monoclonal antibodies. Methods Mol Biol. 2013;988:31-43. doi: 10.1007/978-1-62703-327-5_3.
132. Dashivets T, Thomann M, Rueger P, et al. Multi-angle effector function analysis of human monoclonal IgG glycovariants. PloS One. 2015;10(12):e0143520.
133. Lee LY, Lin CH, Fanayan S, et al. Differential site accessibility mechanistically explains subcellular-specific N-glycosylation determinants. Front Immunol. 2014;5(404).
134. Thaysen-Andersen M, Packer NH. Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching. Glycobiology. 2012;22(11):1440–1452.
135. S-I. N, Itoh A, Nakakita Y, et al. Cooperative interactions of oligosaccharide and peptide moieties of a glycopeptide derived from IgE with galectin-9. J Biol Chem. 2016;291(2):968–979.
136. Sterner E, Flanagan N, Gildersleeve JC. Perspectives on anti-glycan antibodies gleaned from development of a community resource database. ACS Chem Biol. 2016;11(7):1773–1783.
137. Gala FA, Morrison SL. The role of constant region carbohydrate in the assembly and secretion of human IgD and IgA1. J Biol Chem. 2002;277(32):29005–29011.
138. Chuang PD, Morrison SL. Elimination of N-linked glycosylation sites from the human IgA1 constant region: effects on structure and function. J Immunol. 1997;158(2):724–732.
139. Mattu TS, Pleass RJ, Willis AC, et al. The glycosylation and structure of human serum IgA1, Fab, and Fc Regions and the role of N-glycosylation on Fcα receptor interactions. J Biol Chem. 1998;273(4):2260–2272.
140. Carayannopoulos L, Max EE, Capra JD. Recombinant human IgA expressed in insect cells. Proc Natl Acad Sci U S A. 1994;91(18):8348–8352.
141. Ubelhart R, Bach MP, Eschbach C, et al. N-linked glycosylation selectively regulates autonomous precursor BCR function. Nat Immunol. 2010;11(8):759–765.
142. Wright JF, Shulman MJ, Isenman DE, et al. C1 binding by mouse IgM. The effect of abnormal glycosylation at position 402 resulting from a serine to asparagine exchange at residue 406 of the mu-chain. J Biol Chem. 1990;265(18):10506–10513.
143. Muraoka S, Shulman MJ. Structural requirements for IgM assembly and cytolytic activity. Effects of mutations in the oligosaccharide acceptor site at Asn402. J Immunol. 1989;142(2):695–701.
144. Cals MM, Guenzi S, Carelli S, et al. IgM polymerization inhibits the Golgi-mediated processing of the mu-chain carboxy-terminal glycans. Mol Immunol. 1996;33(1):15–24.
145. Shade KT, Platzer B, Washburn N, et al. A single glycan on IgE is indispensable for initiation of anaphylaxis. J Exp Med. 2015;212(4):457–467.
146. Moh ES, Lin CH, Thaysen-Andersen M, et al. Site-specific N-glycosylation of recombinant pentameric and hexameric human IgM. J Am Soc Mass Spectrom. 2016;27(7):1143–1155.
147. Pabst M, Kuster SK, Wahl F, et al. A microarray-matrix-assisted laser desorption/ionization-mass spectrometry approach for site-specific protein N-glycosylation analysis, as demonstrated for human serum immunoglobulin M (IgM). Molecul Cell Prot. 2015;14(6):1645–1656.
148. Loos A, Gruber C, Altmann F, et al. Expression and glycoengineering of functionally active heteromultimeric IgM in plants. Proc Natl Acad Sci U S A. 2014;111(17):6263–6268.
149. Plomp R, Hensbergen PJ, Rombouts Y, et al. Site-specific N-glycosylation analysis of human immunoglobulin E. J Proteome Res. 2013;13(2):536–546.
150. Amin AR, Tamma SM, Oppenheim JD, et al. Specificity of the murine IgD receptor on T cells is for N-linked glycans on IgD molecules. Proc Natl Acad Sci U S A. 1991;88(20):9238–9242.
151. Hoffmann M, Marx K, Reichl U, et al. Site-specific O-glycosylation analysis of human blood plasma proteins. Molecul Cell Prot. 2016;15(2):624–641.

1.6 THESIS AIMS

Overall aim

The overall aim of this thesis was two-fold. Firstly, to improve the characterisation and automation of the analysis of glycans released from proteins (glycomics) for the purpose of increasing the accuracy and throughput of mass spectrometric glyco-analysis. Secondly, to use glycomics and proteomics to characterise the toll-like receptor 4 (TLR4) glycoprotein activation which is one source of inflammation for the central nervous system.

Specific aims

1. Optimise the chromatographic and fragmentation parameters for the mass spectrometric analysis of glycans released from proteins using porous graphitised carbon liquid chromatography electrospray mass spectrometry (PGC-LC-ESI-MS/MS)
2. Develop a foundation for automating mass spectrometric glycan analysis by developing widely available data analysis tools using open-source and vendor-neutral software
3. Apply glycomics analysis using PGC-LC-ESI-MS/MS to investigate protein glycosylation changes following TLR4 activation in LPS stimulated human cell lines and in a mouse model of neuropathic pain.
4. Use proteomics to understand how the TLR4 protein complex with MD-2 and CD14 responds to TLR4 activation in LPS stimulated human cell lines.

CHAPTER 2 - MATERIALS AND METHODS

2.1 MATERIALS

All chemicals were sourced from Sigma Aldrich (Sydney, Australia) unless otherwise specified. Peptide:N-glycosidase F (PNGase F, product#:V4831) was obtained from Promega (Sydney, Australia). Bovine ribonuclease B (product#:R7884), Porcine gastric mucin (product#:M1778), Human IgA (product#:I1010), Bovine lactoferrin (product#:L9507), Human lactoferrin (product#:L0520), Bovine fetuin (product#:F3385) and Human IgG (product#:I4506) were sourced from Sigma Aldrich (Sydney, Australia). Human neutrophil elastase (product#:342-40) was sourced from LeeBio (Maryland Heights, USA). Fungal cellobiohydrolase I was isolated as previously described[23]. All other chemicals were sequencing grade.

Recombinant human TLR4 and MD-2 expressed from HEK293 cells were purchased from Origene (Rockville, MD, U.S.A). Recombinant human MD-2 expressed in *E. coli* was purchased from MyBioSource (San Diego, CA, U.S.A). Heavy phenylalanine labelled, cysteine alkylated MD-2 peptide GSDDDYSFCR was purchased from Cambridge isotope laboratories (Cambridge, MA, U.S.A.).

2.2 CELL LINE CULTURING

The U87MG (ATCC® HTB-14™), J774, BV2, PC12, HEK293, HEK293 +TLR4+MD-2+CD14 cell lines were cultured in a T-75 flask with 10mL of Dulbecco's Modified Eagle's Medium supplemented with 10% (v/v) fetal bovine serum, 50U/mL penicillin, 50ug/mL streptomycin and a final concentration of 2mM L-glutamine. For the stably transfected HEK293FT cell line, the culture medium was also supplemented with 10ug/ml blasticidin and 50ug/ml Hygromycin B Gold to maintain only stably transfected cells. The PC12 cell line were cultured with a slightly different medium in a T-75 flask containing with 10mL of Dulbecco's Modified Eagle's Medium supplemented with 5% (v/v) fetal bovine serum, 5% (v/v) horse serum, 50U/mL penicillin, 50ug/mL streptomycin and a final concentration of 2mM L-glutamine.

After the adherent cells reached 70% confluency, the growth medium was collected and adherent cells scraped and collected. The cells were washed with phosphate buffered saline, centrifuged (500g for 10 minutes) and then the supernatant was removed, for a total of three washes.

For serum-free media experiments, EX-CELL 293 (Sigma Aldrich, Sydney, Australia) was reconstituted as specified by the manufacturer and used with earlier culture methods with slight modification. Once cells reached 70% confluency in the standard culture media containing serum, the culture medium was removed and cells were washed three times with phosphate buffered saline (PBS) and the PBS was discarded. 10mL of the reconstituted EX-CELL 293

medium was then added to the cells and they were left to acclimatise for 24 hrs before performing time-point experiments.

For LPS and LPS-RS treatment experiments, 50ng/mL of LPS or LPS-RS was added to serum-free media 24 hours after cell conditioning with cells at 70% confluency. Adherent cell viability was assessed with the Muse Cell Analyser with the Count and Viability assay kit (Millipore, Australia) following manufacturer's instructions. Cell analyser performance was calibrated before sample analysis using the manufacturer's standard bead solution as the calibrant.

2.3 MOUSE MODEL FOR NEUROPATHIC PAIN

Pathogen-free adult male Balb/c mice (300–350 g; University of Adelaide, Laboratory Animal Services, Waite Campus, Urrbrae, Australia) were utilized in all experiments in this study. Rats were housed in temperature (18–21◦C) and light controlled (12h light/dark cycle; lights on at 07:00h) rooms where standard rodent food and water was available *ad libitum*. Preceding experimentation, rats were habituated to the animal holding care facility for 1 week, followed by 1 week of extensive experimenter handling and acclimatization to the von Frey testing apparatus in order to reduce successive handling stress. All procedures were approved by the Animal Ethics Committee of the University of Adelaide and were performed in accordance with the NHMRC Australian code of practice for the care and use of animals for scientific purposes and adhered to the guidelines of the Committee for Research and Ethical Issues of the IASP.

This study utilized a novel graded sciatic nerve injury model of allodynia, a modified CCI model in the mouse where 4 chromic gut sutures were placed around the sciatic nerve, to develop graded behavioural allodynia (varying degrees of allodynia) as described in detail previously(Graceetal.,2010). Animals were split into three treatment groups, no treatment, sham surgery and CCI surgery with four, five and five animals, respectively.

The von Frey Test was used to investigate mechanical allodynia using phasic stimulation of on Frey filaments across a range of thresholds. Briefly, mice were subjected to 10 stimulations with 6 calibrated von Frey filaments (0.04, 0.07, 0.16, 0.4, 0.6 & 1; grams of force). Von Frey filaments were applied for 1 s at 1 s intervals in random force assignment at each test session. In order to avoid sensitization a 10 min break was given between each set of simulations. The response frequency at each von Frey filament and behavioural responses were recorded as the average number of responses out of 10 for each von Frey filament. Throughout the study, testing was performed blind with regard to group assignment.

On day 22 post-surgery, mice were deeply anesthetised using an i.p. injection of 60mg/kg Lethabarb. The animals were then transcardially perfused using cold 0.9% saline and the lumbar enlargement (L3-L5) of the spinal cord, quickly removed and snap frozen and stored at -80°C until further

assessment. A similar protocol was used for the brain sectioning with the brain quickly removed, snap frozen, PAG and RVM containing regions sliced in 50 µm sections in a cryostat and stored at 80°C until further assessment.

2.4 SAMPLE PREPARATION

2.4.1 RT-PCR and visualisation

Cell pellets (1x106) were chemically and mechanically lysed using lysis buffer [80] (200mM Tris-HCl (pH8), 25mM EDTA, 300mM NaCl and 2% (w/v) SDS) and a syringe equipped with a 20 gauge needle, respectively. The RNA was then precipitated using ethanol and purified using a spin column (RNeasy MinElute Cleanup Kit, Qiagen, Australia), following instructions from the manufacturer. The purified RNA was then quantified with Nanodrop (Thermo Scientific, Australia). The primers used in this study were designed with the NCBI Primer-Blast tool using the plasmid sequences provided by the stably transfected HEK293 cell line vendor. The primers were synthesised by Integrated DNA technologies (Melbourne, Australia).

Reverse transcription reactions were performed in 20µL containing 4µL of SSIV buffer, 1µL of 100 mM dithiothreitol, 1µL of 10mM deoxynucleoside triphosphate (dNTP), 1µL of RNase inhibitor (RNaseOUT, GIBCO), 1 µg of RNA template, 1 µL of 50uM reverse primer, 1µL of reverse transcriptase enzyme (Superscript IV RT, Thermo Scientific, Australia) and Milli-Q water to 20µL. Annealing of the primers to template RNA was performed at 65°C for 5 min followed by incubation on ice. For the reverse transcriptase reaction, samples were incubated at 55°C for 10 min, followed by inactivation of reaction by incubating at 80°C for 10 min.

PCR of the resultant cDNA was performed in 25µL containing 12.5µL of AmpliTaq buffer, 1µL of 5uM forward primer, 1µL of 5uM reverse primer, 1µL/1000ng of cDNA and 9.5µL of Milli-Q water. Activation of the Taq enzyme was performed at 95°C for 10 min, followed by 40 cycle repeats of the following PCR cycle timetable: 95°C for 30 seconds (denaturation), 60°C for 30 seconds (annealing), 72°C for 60 seconds (extension) and 72°C for 7 min (final extension). Once the PCR timetable was completed, samples were held at 4°C.

Visualisation of the RT-PCR products was performed using a 2% agarose TAE gel made from 2g of Ultra-pure agarose, 100mL 1X TAE buffer and 10µL of SYBR Safe Dye. The DNA ladder used was 1kb Plus DNA Ladder (Thermo Scientific, Australia) and samples were mixed with TrackIt Cyan/Orange Loading Buffer (Thermo Scientific, Australia). The electrophoretic solution was 1x TAE buffer and the gel was run at 80V for 30 min, followed by visualisation on a gel imager (G-Box, SynGene).

2.4.2 Protein extraction from tissue

Tissue samples were kept at -80°C until analysis. Mechanical lysis was performed using probe sonication (0.75mL of lysis buffer was added to each tube and probe sonicated twice for 5 seconds and 50% amplitude). This was followed by chemical lysis. (The solutions were transferred to falcon tubes filled with chloroform and methanol. The original sample tubes were washed with 0.25mL H_2O then transferred to the now chloroform, methanol, water solutions. The solutions were mixed for 4 hours to ensure complete disintegration of the tissue into corresponding non-polar, protein and polar layers. The samples were spun down and the protein precipitated interphase was removed and placed in a separate tube. The samples were then dried with centrifugal evaporation and kept at -80°C. Prior to glycomic analysis, samples were resuspended in 8µL of 8M urea. The entirety of each sample was used for *N*-glycan analysis.

2.4.3 Subcellular protein fractionation

Chemical subcellular fractionation was performed as per Ramsby *et al*[24]. Briefly, approximately $1x10^7$ washed cells were resuspended in 5 volumes of ice-cold digitonin buffer (0.01% digitonin (v/v), 10mM PIPES, pH 6.8, 300mM EDTA, 1.2mM PMSF) and incubated on ice with shaking at 300rpm for 10 min. The samples were then centrifuged 500g for 10 min and the supernatant, representing the cytosolic fraction, was flash frozen and stored at -80°C. The pellet was then resuspended in the same volume, as previously used with the digitonin buffer, of ice-cold Triton extraction buffer (0.5% Triton X-100 (v/v), 10mM PIPES, pH 7.4, 300mM sucrose, 100mM NaCl, 3mM MgCl2, 3mM EDTA, 1.2mM PMSF) and incubated on ice with agitation at 300rpm for 30 min. The samples were then centrifuged 5000g for 10 min, the supernatant, representing the membrane fraction, and the pellet, representing the nuclear/cytoskeletal fraction, were separately flash frozen and stored at -80°C.

Following chemical subcellular fractionation, proteins were extracted from the various chemical lysis buffers using a chloroform:methanol:water extraction as previously used for the total cell lysate protein extraction. In all cases, the precipitated proteins were re-solubilized in 4M urea and protein yield quantified with the Bradford Protein assay[25].

2.4.4 Cell lysis, protein precipitation and quantification

To obtain a total cell lysate, washed cells were lysed and cellular protein precipitated using a chloroform:methanol:water extraction[26] (10:10:3, by volume) by vigorous agitation (1000rpm) for 30 min at 4°C. For secreted proteins, the spent culture media was added to ice cold acetone in a ratio of 1:9 and left overnight at -80^{o}C. For both protein precipitation steps, samples were then centrifuged at 5000g for 5 min and protein precipitate was removed and stored at -80°C. Prior to performing protein quantitation, all samples were re-solubilized, or diluted, to a final concentration of 4M urea and protein yield quantified with the Bradford Protein assay[14].

2.4.5 Reduction, alkylation and SDS-PAGE analysis of proteins

Preparation of protein extracts for mass-spectrometric analysis was performed as per Shevchenko *et al* with some modifications[27]. Approximately 10µg of extracted protein was reduced using dithiothreitol and alkylated using 2-iodoacetamide, followed by mixing with LDS sample buffer (NuPAGE, Thermo Scientific, Australia) before loading into the gel. After the molecular weight ladder was loaded (Novex Sharp Pre-stained Protein Standard, Thermo Scientific, Australia), separation was performed with a 4-12% Bisacrylamide-Tris sodium dodecyl sulphate gel (Invitrogen, Melbourne, Australia) with 1X MOPS buffer at 220V for 30 min. Extracted proteins were visualised in the gel with Coomassie blue staining. Gel bands corresponding to MD-2 and TLR4 (approximately 22kDa and 110kDa, respectively) were excised.

2.4.6 Trypsin digestion and C18 Cleanup

For in-solution trypsin digestion, reduced and alkylated proteins were digested for 16 hours by a trypsin/Lys-C mixture (Promega, Australia) with approximately 1:100 ratio of enzyme to protein. Enzymatic activity was quenched following addition of formic acid. For in-gel trypsin digestion, reduced and alkylated proteins in polyacrylamide gel pieces were cut to roughly 1mm3 cubes and then digested as described by Shevchenko *et al*[27]. Following digestion, peptides were extracted from gel pieces.

Both peptide gel extracts and the result of in-solution proteolytic digestion were dried by vacuum centrifugation. Following this, peptides were enriched and desalted using C18 Omix tips (Agilent, Australia) per the manufacturer's instructions. The subsequent eluate was dried by vacuum centrifugation and resuspended in Milli-Q water with 0.1% formic acid (v/v), centrifuged at 16,000g and the liquid was transferred to autosampler compatible vials (Waters, Australia) for LC-MS analysis.

2.4.7 N-glycan and O-glycan release

N- and *O*-glycans were released from immobilised protein samples as described by Jensen *et al*[15] unless otherwise specified. Briefly, 10µg of the protein samples were spotted on polyvinylidene difluoride (PVDF) membranes (Millipore, Sydney, Australia) and stained with Direct Blue (Sigma-Aldrich, Sydney, Australia). The membrane spots were excised and washed in separate wells in a flat bottom polypropylene 96-well plate (Corning Incorporated, NY). *N*-glycans were released from the membrane-bound protein using 1U PNGase F (Promega, Sydney, Australia) with overnight incubation at 37°C. Following *N*-glycan removal, 500 mM $NaBH_4$ in a 50 mM KOH solution was added to the membrane spots for 16 h to release reduced *O*-linked glycans by reductive β-elimination.

Released *N*-glycans were reduced with 1 M $NaBH_4$ in a 50 mM KOH solution for 3 h at 50˚C, after which the reaction was neutralized by adding equimolar glacial acetic acid. Both *N*-glycans and *O*-glycans were desalted and enriched offline using AG 50W-X8 (Bio-Rad) strong cation exchange

followed by PGC solid phase enrichment micro-columns (Grace, Columbia, MD, USA) prior to analysis.

In-gel glycan release was performed as described by Royle *et al*[28]. Briefly, excised gel pieces in an Eppendorf tube were cleaned with successive water/acetonitrile washes and dried with vacuum centrifugation followed by overnight digestion at 37°C with 0.5 units of PNGase-F in 20mM ammonium bicarbonate. Released *N*-glycans were collected by washing the gel pieces with water and acetonitrile. Samples were then reduced, desalted and enriched as described by Jensen *et al*[29], followed by mass-spectrometric analysis.

2.4.8 Glycan permethylation

Permethylated glycan samples were prepared following the above steps, following by solid-phase spin-column derivatisation and subsequent clean-up as described by Mechref *et al*[30]. Briefly, dried samples from PGC sample preparation were permethylated by loading the sample to a dry NaOH bead bed within a MacroSpinColumn (Harvard Apparatus, MA, USA). 15µL of methyl iodide was added to the sample in the NaOH bed. The samples were then eluted from the column with acetonitrile followed by drying the sample in a SpeedVac.

Dried permethylated glycans were then desalted by liquid-liquid extraction in chloroform and water and then enriched offline with C18 Omix tips (Agilent, MA, USA).

2.5 DATA ACQUISITION

2.5.1 Direct infusion parameters

Direct infusion was performed at 5ul/min using a 250µL needle, until at least 500 scans were collected. The direct infusion was monitored with a Linear Trap Quadrupole mass spectrometer (Thermo Scientific LTQ Velos Pro) according to these MS conditions: *m/z* 570–2000, 1 microscan, *m/z* 0.35 resolution (FWHM), 1×10^4 automatic gain control (AGC) and 100 ms accumulation time. Spectral data was acquired in profile mode.

2.5.2 PGC-LC-ESI-MS/MS glycan analysis

PGC-LC/ESI-MS/MS experiments were performed on an UltiMate3000 high performance liquid chromatography (HPLC) system (Dionex, Sunnyvale, CA, USA) interfaced with a Linear Trap Quadrupole (LTQ) Velos Pro ion trap (Thermo Scientific, San Jose, CA, USA) unless stated otherwise. Some PGC-LC/ESI-MS/MS experiments were also performed on an 3D ion trap using an Agilent 1100 capillary LC system (Agilent Technologies, Santa Clara, CA) interfaced with an Agilent 6330 LC-MSD 3D Trap XCT ultra.

Separations on both instruments were performed using a PGC LC column (3 µm, 100 mm x 0.18 mm, Hypercarb, Thermo Scientific) maintained at room temperature and at 50°C for the Agilent and Dionex LC systems, respectively. 10 mM ammonium bicarbonate aqueous solution (solvent A) and 10 mM ammonium bicarbonate aqueous solution with 45% acetonitrile (v/v) (solvent B) were used as mobile phases. The flow rate was 2 µL/min and 4 µL/min for the Agilent and Dionex LC systems, respectively. The same gradient was used for both systems with the following linear gradient program: 0 min, 2% B; linear increase up to 35% B for 53 min; linear increase up to 100% B for 20 min; held constant for 5 min; and then equilibrated at 2% B for 5 min before next injection - giving a total LC run time of 83 min. The ESI-MS2 analysis was operated in negative ion mode with source voltage at -3.2 kV for both instruments.

For the Thermo Scientific linear ion trap (LTQ), glycans were analysed according to these MS conditions: *m/z* 580–2000, 3 microscans, *m/z* 0.25 resolution (FWHM), 5×10^4 automatic gain control (AGC) and 50 ms accumulation time and MS/MS conditions: *m/z* 0.35 resolution (FWHM); 2×10^4 AGC, 300 ms accumulation time, 2 *m/z* window and top five data-dependent acquisition. During MS/MS scans, RE-CID fragmentation was used with helium as the collision cell gas. RE-CID subjected ions to 33% normalized collision energy (NCE) with an activation Q of 0.250 and an activation time of 10 milliseconds (ms). For ion trap HCD (Sup Fig 7), nitrogen was used as the collision gas. HCD subjected ions to 30-37.5% NCE with a default charge state of 2 and an activation time of 2ms. Spectral data were acquired in profile mode.

For the Agilent 3D ion trap, glycans were analysed according to these MS conditions: *m/z* 350–2200, 5 microscans, *m/z* 0.13 resolution (FWHM), 8×10^4 ion current control (ICC) and 200 ms accumulation time and MS/MS conditions: *m/z* 0.13 resolution (FWHM), 8×10^4 ICC, 200 ms accumulation time, 4 *m/z* window and top three data-dependent acquisition. RE-CID acquisition, using helium as the collision cell gas, was performed for characterization of the glycan ions detected in MS/MS scans. Fragmentation amplitude was set to 1 V with “Smart Frag” enabled ramping from 30% to 200% of the fragmentation amplitude for RE-CID with an activation time of 40 ms. All spectral data were acquired in profile mode.

2.5.3 C18-LC-ESI-MS/MS permethylated glycan analysis

C18-LC/ESI-MS/MS experiments were performed on a 1200 series nanoflow HPLC-Chip nanoESI HPLC with a HPLC-Chip consisting of a 360nl trapping column and a 150 mm x 75 um analytical column, both with Polaris C18-A 3 um resin, interfaced with a 6550 iFunnel Quadrupole Time-of-Flight (QTOF) (All from Agilent, Santa Clara, CA). Separations were performed using the C18 Chip HPLC column with an aqueous buffer composed of 0.1% formic acid (v/v), 1% acetonitrile (v/v) (solvent A) and 99% acetonitrile (v/v) with 0.1% formic acid (v/v) (solvent B) used as mobile phases. Following sample injection, permethylated glycans were trapped on the C18 Chip trap with washing at 2ul/min with

solvent A. Permethylated glycans were then separated on the analytical column at a flow rate of 0.3ul/min using a gradient as: 0-5min, 10% B; linear increase over 35 minutes to 65% B; linear increase over 10 minutes to 95% B; held at 95% B for 5 minutes; linear decrease over 5 minutes to 10% B and then held at 10% B for 5 minutes, for a total of 65 minutes run time.

The ESI-MSn analysis was operated in positive ion mode with source voltage at +1.45. Glycans were analysed according to these MS and MS/MS conditions: *m/z* 295-1700 with 15,000 resolving power @ *m/z* 200 (FWHM). The top 20 MS peaks in the range between *m/z* 295–1,700 were selected for MS/MS fragmentation, and a custom ramped collisional energy (CE) table was used to determine the CE, based upon precursor charge. Nitrogen was used as the collisional gas.

2.5.4 C18-LC-ESI-MS/MS peptide analysis

QTOF-based proteomic data was acquired on a 5600 TripleTOF+ mass spectrometer (SCIEX, Framingham, MA, USA) coupled with an Eksigent Ultra-nanoLC-1D system (Eksigent, Dublin, CA, USA) Ten microliter peptide samples containing 1µg of peptide (as determined by the Bradford assay[25]) were injected onto a peptide trap (Bruker peptide Captrap) for preconcentration and desalted at 10 µl/min for 5 min with 0.1% formic acid (v/v) and 2% acetonitrile (v/v).

After desalting, the peptide trap was switched in-line with an in-house packed analytical column (75 µm × 10 cm of solid core Halo C18, 160 Å, 2.7 µm media (Bruker, Bruker Manning Park Billerica, MA, USA) and fused silica PicoTip emitter (New Objective, Woburn, MA, USA). Peptides were eluted and separated from the column using the buffer B (99.9% (v/v) acetonitrile, 0.1% (v/v) formic acid) gradient starting from 2% and increasing to 30% over 60 min at a flow rate of 600 nl per minute. A TOFMS survey scan was acquired at *m/z* 100–1500 with 0.1 s accumulation time, and the precursors from the inclusion list were consecutively isolated for product ion scans. Product ion spectra were accumulated for 0.5 milliseconds in the mass range *m/z* 100–1500 with rolling collision energy ((0.0625 × *m/z* – 3) for z = 2, (0.0625 × *m/z* – 4) for z = 3, and (0.0625 × *m/z* – 5) for z = 4).

Q-Orbitrap-based proteomic data was acquired on a Q-Exactive hybrid quadrupole Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA) with the same settings as previously described except for the MS1 and MS2 parameters. Peptides were analysed according to these MS1 conditions: 350–2000 *m/z*, 1 microscan, 70,000 resolution, 5×10^4 AGC and 100 ms injection time and MS2 conditions: 35,000 resolution, 2×10^5 AGC, 500 ms injection time, 2 *m/z* isolation window, 30 normalized collision energy (NCE). Precursor ions were selected for MS/MS using a dependent scan for the ions specified in the inclusion list. Similar to C18-LC-ESI-QTOF-based proteomic analysis, peptides were loaded onto a 10 cm x 75µm column packed with C18 2.7µm resin (Advanced Materials Technology, Wilmington, DE, USA) at 600 bar. Samples were analysed using an Easy-nLC 1000 UHPLC system (Thermo Scientific, San Jose, CA, USA) configured to a Q Exactive hybrid

quadrupole Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA). A 40 minute gradient was run at 300 nL/min using 0.1% formic acid (v/v) (Buffer A) and 99.9% acetonitrile (v/v) with 0.1% formic acid (v/v) (Buffer B). The gradient was 1% Buffer B to 50% Buffer B across 30 min, then to 85% Buffer B for 2 min, then held at 85% Buffer B for a total run time of 40 min.

The column eluent was directly ionised at 1.6kV with a capillary temperature of 250°C. Peptides were analysed according to these MS1 conditions: 350–2000 *m/z*, 1 microscan, 35,000 resolution, 1×10^6 AGC and 120 ms injection time and MS2 conditions: 17,500 resolution, 2×10^5 AGC, 60 ms injection time, 1.4 *m/z* unit window, top 10 data-dependent acquisition and 30% normalized collision energy (NCE).

2.6 DATA ANALYSIS

2.6.1 Manual glycan identification

Lists of ions subjected to RE-CID MS/MS (-) from relevant spectra of each sample were extracted using ESI-compass v1.3 Bruker Daltonic Software (Bruker DALTONIK GmbH, Bremen, Germany) and Xcalibur v3.0.63 (Thermo Scientific, Australia) for the Agilent 3D ion trap and Thermo LTQ, respectively. After removal of common contaminants, the extracted monoisotopic precursor masses potentially corresponding to glycans were searched against GlycoMod (http://www.expasy.ch/tools/glycomod) to identify putative monosaccharide compositions.

Interpretation and validation of glycan identities were based on the existence of A-, B-, C-, X-, Y- and Z- product ions consistently found across the majority of MS/MS scans over the elution times of the respective precursor ions. GlycoWorkBench v2.1 (available from https://code.google.com/archive/p/glycoworkbench/) was used for most glycan product ion annotation with an in-house monosaccharide product ion database used to cover fragment gaps. A supplementary GlycoWorkBench file is included with this thesis, containing annotated MS2 spectra representing the glycan structures quantified. Product ions are assumed to be singly charged if corresponding *m/z* value was observed to be greater than the precursor and/or no isotopic distribution is observed which assigns a product ion charge state equal or lower to the precursor. RawMeat v2.1 (Vast Scientific) was used to analyse Thermo Scientific raw files for run statistics regarding instrument duty cycle.

2.6.2 Glycan quantitation using Skyline

Vendor-specific raw files were converted to vendor-neutral mzML file formats sing Proteowizard v 3.0.11676 (http://proteowizard.sourceforge.net/). These .mzML files were then converted to a human readable format using msaccess (as part of the Proteowizard suite, http://proteowizard.sourceforge.net/) and read using Excel (Microsoft Office 2010). Skyline (64-bit) v3.6.0.10493 (https://skyline.ms/) was used for all analyses at both the MS and MS/MS level.

Full scan settings: For the MS-level filtering, isotope peaks were included by count, only filtering for the monoisotopic peak to be used for precursor ion peak area calculation. The precursor mass analyser was set to a quadrupole ion trap (QIT) with *m/z* 0.5 and *m/z* 0.35 resolution (FWHM) for the Agilent 3D ion-trap and the Thermo linear ion-trap (LTQ), respectively. High selectivity extraction was used and no retention time filtering was performed. MS/MS filtering was used with acquisition method set to "targeted" with the product mass analyser set to QIT with *m/z* 0.5 and *m/z* 0.35 resolution (FWHM) for the Agilent 3D ion-trap and the Thermo LTQ ion-trap, respectively. For Agilent and Thermo QITs, precursor ions were allowed to be in the interval *m/z* 50-2200 as the maximum with a method match tolerance of *m/z* 0.6.

The precursor ion targets for Skyline analysis were used from a previously developed list of confirmed glycan composition, and included information of 1) the glycan class (paucimannose, high mannose, complex, hybrid), 2) confirmed monosaccharide composition/isomer elution order 3) experimental precursor ion *m/z*, 4) experimental product ion *m/z*, 5) precursor ion charge state and 6) product ion charge state.

Identification of glycan isomer discriminators was performed by manual analysis of individual MS/MS scans for each glycan isomer in the standards run on the Thermo LTQ. Product ions were deemed to be discriminators by two criteria: 1) presence in majority (>80%) of MS/MS scans for the glycan isomer to be discriminated and 2) detection at greater abundance within the spectrum arising from the targeted glycan isomer compared to all other isomers.

Similar to the MS-level peak picking, the Skyline automatic peak picking algorithm was used but only for manual peak picking of glycans not subjected to MS/MS. The integrated peak areas were exported as per MS-level peak picking. Integrated peak areas were exported using a custom report format made for glycan analysis.

2.6.3 Peptide identification and quantitation

FASTA files for TLR4 and MD-2 isoform 1 protein sequences from UniProt (O00206 and Q9Y6Y9, respectively) were used with Skyline (version 4.1) to build the targeted proteomics inclusion list. Skyline parameters for building this list were: Enzyme: Trypsin, maximum missed cleavages: 1, minimum peptide length: 4, maximum peptide length: 25 and modifications: carbamidomethylation of cysteine. Populating the theoretical peptide product ion list, only singly charged product ions, arising from *y* and *b* -type cleavages, were considered

Proteome Discoverer v2.2.0.388 (Thermo Scientific, CA, USA) was used for protein identification from shotgun proteomic experiments. Spectrum selection was *m/z* 350-2000. SequestHT was used as the search engine using the SwissProt non-redundant protein database with the following search parameters: precursor tolerance 10ppm, product tolerance 0.02 Da, 2 maximum missed cleavages

and peptide length 6-14 amino acids. The following dynamic modifications were included: Oxidation (methionine) and acetylation (N-terminus). Carbamidomethylation (Cysteine) was a static modification. False discovery rate was 5% at the peptide level.

Byonic (Protein Metrics, CA, USA)[31] parameters were set as follows: Digestion parameters: cleavage site: KR, Cleavage side: C-terminal, Digestion specificity: Fully specific, Missed cleavages: 2. Instrument parameters: Precursor mass tolerance: 30 ppm, Fragmentation type: QTOF/HCD, Fragment mass tolerance: 50 ppm. Modifications: carbamidomethyl of cysteine (Fixed), oxidation of methionine (Common2), deamidation of asparagine (Common1). Glycans: High mannose/Paucimannose list and *N*-glycan 309 mammalian list used. Fasta files used for searching were the TLR4 and MD-2 protein sequences from UniProt (O00206 and Q9Y6Y9, respectively).

Searching of PRIDE (https://www.ebi.ac.uk/pride/archive/) was performed using UniProt accessions for human TLR4 and MD-2 isoform A (O00206 and Q9Y6Y9), last accessed December, 2017. Skyline was used for validation of identifications with the following settings:

Full scan settings: For the MS-level filtering, three isotopic peaks were included by count. The precursor mass analyser was set to an Orbitrap with 60,000 resolution (FWHM at *m/z* 200). High selectivity extraction was not used and no retention time filtering was performed. MS/MS filtering was used with acquisition method set to “targeted” with the product mass analyser set to Orbitrap with 35,000 resolution (FWHM at *m/z* 200). For Agilent and Thermo QITs, precursor ions were allowed to be in the interval *m/z* 50-2000 as the maximum with a method match tolerance of *m/z* 0.055.

Protein Blast (BlastP, https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) was performed using peptide sequences reported in PRIDE datasets for TLR4 or MD-2 detection using accessions for isoform A (O00206 and Q9Y6Y9 against the NCBI non-redundant database with no specific organism. Only protein sequence matches with 100% sequence query cover and identity were considered.

Microsoft Excel 2010 was used for a variety of data analysis steps. Glycan abundance plots were performed as described by Weissgerber TL *et al* [32]. Ratios for log2 glycan abundance heat maps was performed by determining the fold change compared to time point 0, briefly, by dividing the relative abundance of a glycan at the time point of interest by its relative abundance at time point 0. On these values, a Log2 transformation is performed and displayed as a heatmap based on their values compared to time point 0. Duty cycle visualisation and calculation was performed values from the instrument method and Velos Pro Instrument Setup Help from Xcalibur V2.2.

Microsoft Excel was also used for determination of behavioural measurements, allodynia and hyperalgesia. Allodynia was calculated by the animal behavioural response with von Frey hair

number 1 on day 21 of behavioural testing (Supp Fig 7.1). Hyperalgesia was determined to be the slope of the linear trend line corresponding to von Frey hair number and animal response.

2.6.4 Databases

A variety of glycan databases were used through the course of this thesis. GlycoStore (https://www.glycostore.org/) was used to assist in glycan structure assignment using PGC elution order. Structural determination was also assisted by UniCarbKB (http://www.unicarbkb.org/) in identifying reported glycan structures matching the glycan composition determined using GlycoMod (https://web.expasy.org/glycomod/) and these glycan structures were used as potential matches. The Consortium for Functional Glycomics (http://www.functionalglycomics.org/) publicly available datasets were used to compare to our mouse tissue glycan profiles. GlyTouCan (https://glytoucan.org/) was used to report glycan structures identified as a result of these studies with structures added if not originally in database.

SECTION 1 – METHOD DEVELOPMENT IN THE FIELD OF GLYCOMICS FOR IMPROVED GLYCAN CHARACTERISATION AND QUANTITATION

Formation of glycan ions for detection by mass spectrometry

Mass spectrometry is dependent on the manipulation and measurement of ions in the gas phase, therefore the manner of ion formation is a worthwhile starting point for optimisation. In current glycomics experiments, there are two predominant methods to detect glycans: electrospray ionisation (ESI) and matrix-assisted laser desorption ionisation (MALDI). For ESI and MALDI, both techniques aim to generate glycan ions from solutions or crystals, respectively, for subsequent mass-spectrometric detection.

The type of ionisation has an influence over the ion populations observed. MALDI typically generates singly charged ions; conversely ESI is capable of forming multiply charged ions, [33]. In glycomics experiments, the charge state and adduction of glycans plays a role in their fragmentation and characterisation of their structures using diagnostic ions[34]. As fragmentation techniques typically involve the dissociation of a precursor ion into two or more product ions, a singly charged precursor subjected to collision-induced dissociation (CID) will result in only one product ion possessing the charge (Fig. 3.1). On-line separation of glycan structures can be performed using HPLC with a suitable stationary phase such as porous graphitised carbon (PGC), allowing glycan isomer quantification. As we desired glycan isomer quantitation and used HPLC for the majority of experiments in this dissertation, only ESI will be further discussed.

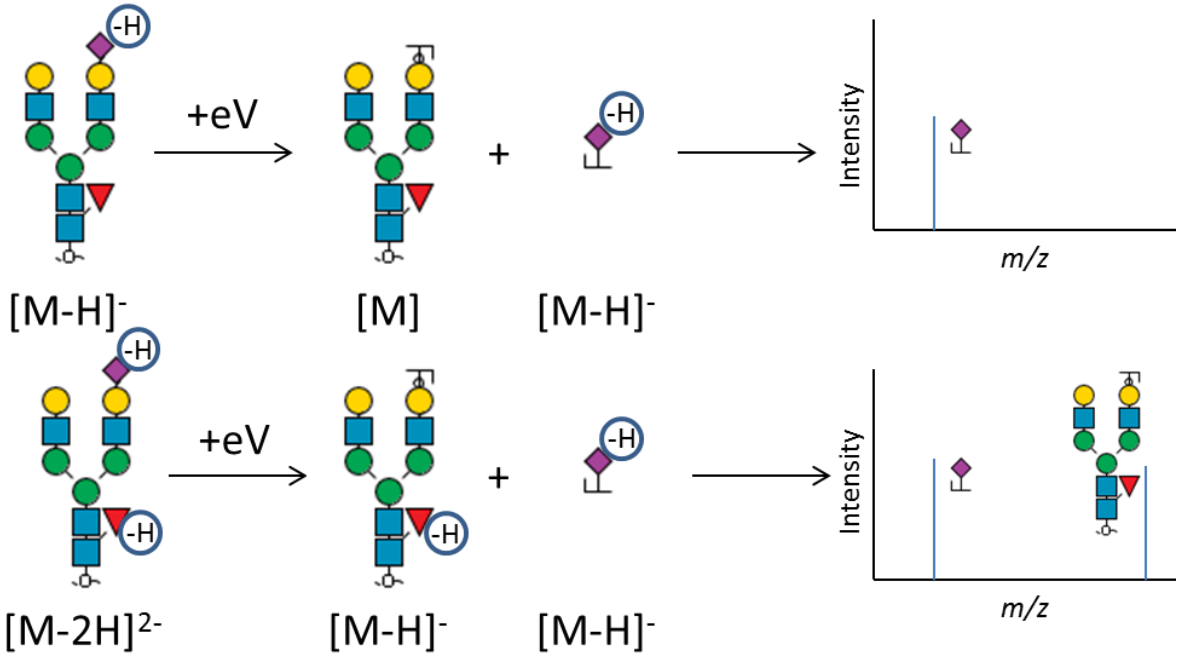


Figure 3.1 Simplified representation of charge state influence in product ion spectra resulting from CID of a glycan precursor

The effectiveness of ESI is dependent on many factors, drastically affecting results obtained from almost identical instruments. As ESI works by neutralising the counter-ion, the buffer composition and flow rate is directly related to its effectiveness. In positive ion mode analyses, this is typically accomplished through utilisation of an acidic buffer (typically containing 0.1%v/v formic acid) which helps promote the formation of positively charged ions. This is contrasted by negative ion mode which requires a solvent capable of creating stable anions, such as 10%v/v 2,2,2-trifluoroethanol/90% methanol which has been identified as one of the best solvents for stable anion generation with negative ion ESI[35]. The needs for ESI also must be balanced with the stationary phase requirements of the liquid chromatographic glycan separation, which is discussed later in this chapter. For non-derivatised mammalian glycan analysis, negative ion polarity analysis is the predominant ion analysis method used due to the presence of sialic acids which have an inherent negative charge.

Ion-trap mass spectrometry

The ion trap mass-analyser is sensitive and versatile, capable of both signal quantitation (using MS1) and characterisation of the molecule responsible for the quantified signal (using MS1 and MS2). While ion-traps have a lower resolution compared to FTICR, Orbitrap or QTOF mass analysers, ion-traps are lower cost and therefore are highly represented in life-science research.

While high-resolution, the ability to discriminate molecules with similar masses, is desirable when dealing with complex mixtures of molecules with similar masses such as peptides, glycans are typically well resolved in the mass dimension (aside from isomers) and do not require high resolution for the vast majority of glycan structures, except for identifying phosphate or sulphate modifications (79.9663 Da and 79.9568 Da, respectively). For analysis of product ions from fragmentation events, high resolution, accurate mass analysis could offer more confident structure elucidation due to its ability to discriminate between isobaric glycan fragments[36].

The development of the modern-day ion-trap began with Wolfgang Paul and co-workers who invented the quadrupole ion-trap[7]. Through subsequent developments and improvements in manufacturing, a two-dimensional QIT[37] (with greater trapping efficiency and ion storage compared to a 3D trap) was invented which was further developed into a dual pressure QIT[38].

Measurement of glycan abundance involves quantification based on area under the curve peak area calculation on extracted ion chromatograms of monoisotopic peaks of glycans. As porous graphitized carbon columns are unable to retain mono- and di-saccharides[29], released glycan masses can vary greatly from 572 Da ($(HexNAc)_2$ $(dHex)_1$) to over 3600 Da ($(Neu5Ac)_4(Hex)_7(HexNAc)_4$ $(dHex)_1$); thus when using negative mode ESI, glycans are detected as multiply charged adducts up to a 4- charge state. A *m/z* 0.25 resolution (FWHM) is therefore required for confident charge state determination

and accurate peak area calculations on the monoisotopic peak of these multiply charged species and this can be provided by modern linear and 3D ion-traps, depending on manufacturer. High resolution, accurate mass analyses can further improve confidence in glycan composition determination[39] compared to lower resolution ion-traps.

Characterisation of glycans using LC-MS

Human disease and development modulate the glycosylation pathway through nucleotide-sugar levels and abundance of enzymes involved in glycosylation (glycosyltransferases and glycosidases). Through characterisation and quantification of glycan substructural features such as glycan fucosylation and sialylation linkages, glycomics can be linked to proteomics and metabolomics, to understand glycosylation at a system level. Furthermore, without complete glycan characterisation, the glycomics community would not be able to identify when newly identified glycan structures are detected and changes are needed in our fundamental knowledge of protein glycosylation.

Characterisation of glycans from an LC-MS experiment can be performed using three discrete forms of information: monoisotopic mass, retention time and MS/MS spectrum of each glycan[40, 41]. While monoisotopic mass is the most widely applicable data for determining the composition of a glycan from an LC-MS experiment, it is not sufficient for discriminating between glycan isomers which share the same exact mass but differ in sequence and/or linkages. Chromatographic retention time can also be insufficient for glycan isomer discrimination if separation with a gradient or stationary phase is incapable of separating these isomers; thereby giving low peak resolution[42]. Depending on fragmentation settings, different glycan isomers can have characteristic product ions, representative of their substructures [43]. Tandem MS may product fragmentation that differentiates co-eluting glycan isomers gives the most informative description of the glycan structure (or structures) present in a peak, particularly for the purposes of isomer discrimination.

The ability to quantitate changes in glycan composition and structure for a given protein or cell is dependent on the methods used. Quantitation of glycans comes in two flavours: absolute quantification, which obtains the absolute concentration of specific glycan structures within a sample, and relative quantitation, which obtains the comparative abundance of specific glycan structures in relation to the glycan structure mixture as a whole[44]. Mass spectrometry-based quantitation of glycans is the second most widely used method for determining changes in glycan structure abundances with LC-UV the most common method[45]. MS-based glycan quantitation can be absolute by using stable isotope labelled standards for each quantified glycan structure but its use as relative quantitation is suitable for identifying glycan structure changes relative to the overall glycan profile[44].

Despite the focus of quantitation for answering biological questions, the implications of suboptimal parameters for LC-ESI-MS are rarely mentioned or discussed in publications.

In this chapter, we present a series of optimisation steps towards developing a vendor-specific method for routine *N*- and *O*- glycan analysis using the Thermofisher Scientific LTQ-Velos Pro. Through assessment of several parameters including ESI spray voltage, MS scan averaging, data acquisition strategy and duty cycle time, we develop a method that serves as a useful starting point for improved glycan characterisation (Fig. 3.1).

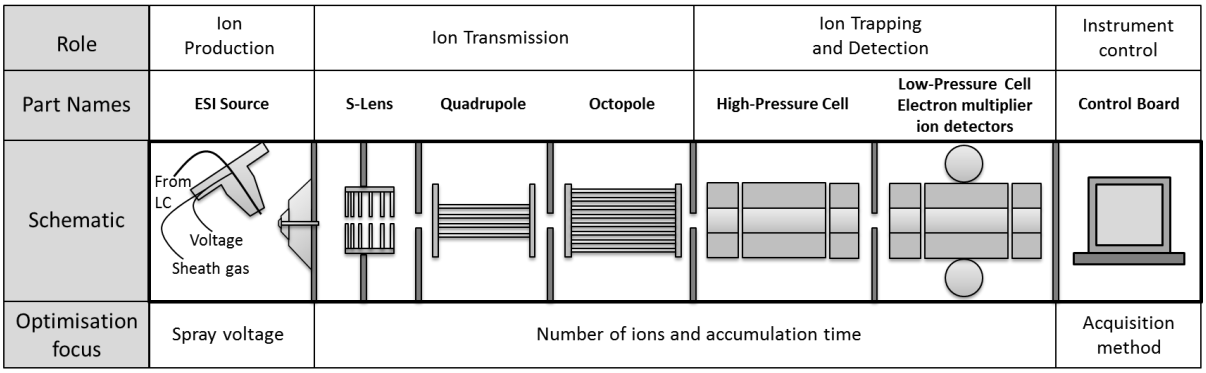


Figure 3.1 Mass spectrometer areas of optimisation for this chapter in context of the Velos Pro mass spectrometer

3.1 INTRODUCTION

Quantitative requirements of a successful glycomics experiment

A successful glycomics experiment, while largely subjective, typically is defined as the characterisation and/or quantitation of glycans released from proteins. As there are no published guidelines on requirements for a successful glycomics experiment, we have taken requirements from both proteomics and metabolomics, as they use similar instrument acquisition methods to quantify and characterise biomolecules. Similar to proteomics and metabolomics, for MS1 or full-scan requirements, glycomics requires a resolution at *m/z* 0.25 or better with equivalent mass accuracy (due to the prevalence of glycans with -4 charge states), 10 MS1 scan points across a peak [46] (to closely approximate the true intensity of the glycan over a chromatographic peak), *m/z* 570-2000 to detect the majority of glycan structures (fucosylated chitobiose core (FM0) and larger) and effective dynamic range to quantify and characterise glycans which can be found in complex mixtures[47]. For reduced, but otherwise native glycan characterisation and quantitation, negative ion mode analysis can be considered essential due to structural rearrangement[48] and low quantitative response for negatively charged glycans that occur in positive ion mode analysis..

The MS2 requirements for successful ion fragmentation have a greater diversity of fragmentation techniques/methods and each technique often has a specific application area. Despite this, we have come up with three requirements for quality MS2 data acquisition in glycomics: reproducible fragmentation (required for spectral library searching and use of diagnostic ions), efficient dissociation of precursor[38] and an isolation width specific for glycans with minimal co-isolation with other ions. Both the MS1 and MS2 requirements inform and shape the data acquisition strategies used for glycomics. Although now the focus of this work, a glycan's isotopic distribution also serves as a way to assign precursor and product ions, through determination of their charge state and corresponding monoisotopic peak[49].

Data acquisition strategies for glycomics

The acquisition of data in a MS-based glycomics experiment is largely dependent on the instrument and the accompanying set acquisition method. The importance of data acquisition strategy has been recently and comprehensively covered for metabolomics [50], which applies equally to glycomics, with a few exceptions based on the nature of glycan biosynthesis. One of the key differences between metabolomics and glycomics is that all mammalian *N*-linked glycans are thought, at present knowledge, to follow a described biosynthetic pathway with multiple divergences resulting in a range of endpoints with certain substructures (e.g. core-fucosylation and the *N*-glycan core).

Metabolomics aims to quantify many metabolites with widely different structural properties, resulting in a wide variety of masses with different atomic elements as part of their molecular composition, thereby requiring high-resolution instruments to successfully determine the molecular formula of a metabolite. In addition to determining the molecular formula, fragmentation can then be used to confirm the metabolite identity but unlike glycomics, metabolomics features curated databases of a wide range of metabolites to be used for MS/MS spectral matching[51]. These key differences can be attributed to the difference in strategy for data acquisition in glycomics.

Untargeted data acquisition

Untargeted data acquisition is a frequently used method for MALDI-based experiments, where the data can be easily re-acquired after some initial data analysis. This then allows for MS/MS or exo-glycosidase experiments to be performed on samples analysed earlier, to further characterise identified glycan structures[52]. Before the value of MS/MS was understood to be valuable for, and thus became routine, for glycan characterisation using LC-ESI-MS, untargeted data analysis was often used[41].

Data dependent acquisition

Data-dependent acquisition (DDA) is a workflow which uses the mass-spectrometer to switch between full-scan MS and MS/MS; with switching based on several defined MS2 criteria including intensity, isotope pattern or neutral loss, just to name a few. This method is very popular in shotgun proteomics, where pre-defined MS1 targets are not used and MS/MS is used to confirm peptide identity. Depending on the instrument, several MS/MS scans can be performed before switching back to full-scan MS, providing characteristic MS/MS spectra for multiple precursors in one duty cycle. This also highlights the main limitation of DDA; that there is a compromise between number of MS/MS scans and MS1 scan frequency. While this is not a particularly worrying limitation for glycomics, due to its relatively low number of theoretically possible glycan structures this has been of particular concern for complex protein mixtures analysed with shotgun proteomics[53].

Variation in protein abundances between replicate samples can lead to missing MS/MS spectra for peptides important for hypothesis verification due to co-eluting analyte interference; with this interference more likely in proteomics due to higher sample complexity compared to glycomics. There are 68 million theoretically possible peptide sequences with lengths less than 8 amino acids[54] whereas only 1.6 million theoretically possible *N*-glycan structures with less than 16 monosaccharide residues[55]. While protein glycosylation also includes *O*-glycosylation, many sample preparation methods utilise sequential *N*- and *O*-glycan releases to reduce the sample complexity and to allow tailored downstream analytical approaches to be applied[29, 42, 56, 57]. In the proteomics field, the response to co-eluting analyte interference was the development of data-

independent acquisition methods such as Sequential Window Acquisition of All Theoretical Spectra (SWATH[58]) and All Ion Fragmentation (AIF[59]).

On an ion trap-based analysis platform, data-dependent analysis can have further depth through the ability to perform MS/MS on product ions generated from MS2, i.e. MS3 and further which is also known as MSn. This technique is frequently used for permethylated glycan analysis, due to the highly informative product ion spectra produced by glycosidic bond cleavages and cross-ring cleavages of these derivatised glycans which can be specific for certain glycan structures. Unfortunately, co-elution of isomeric glycan structures can complicate this data analysis. By subjecting prospective diagnostic product ions to MS3, their identity can be further confirmed. One of the first uses of this approach was in 1998[60] and this strategy continues to be further developed for higher-throughput applications[36].

While ion-mobility has only recently been used for characterisation of glycans [61, 62], the Bruker TIMS-TOF offers an alternate method for acquiring data through the use of trapped ion mobility spectroscopy to enrich and perform high throughput MS/MS scans. This could be a very useful future platform for the high throughput characterisation and quantitation of glycan isomers but there have been no online (LC-MS) applications in the glycomics area yet.

Data independent acquisition

To compete with the data-dependent acquisition strategies, a series of strategies called data-independent acquisition (DIA) have been established to ensure continuous acquisition of MS/MS spectra. Two such techniques include SWATH[58] and AIF[59] but there have been no publications to date using these techniques for the analysis of released glycans, despite its growing popularity in metabolomics and proteomics. We speculate that released glycan mixtures, from a cell lysate proteins for example, have fewer biomolecules compared to a typical cellular proteomics/metabolomics experiment and these techniques typically perform best when there are more ions that are available for MS/MS scans per duty cycle. We also believe that these strategies could be difficult to implement for glycomics, given that glycans share the same substructures, resulting in a lower number of fragment ions specific for one glycan structure over all others, which proteomics typically relies on.

Targeted analysis of glycans

Much like data-independent analysis, targeted analysis of glycans is not frequently published, with few publications detailing targeted methods to detect glycans including SRM and MRM. This topic has been well reviewed by Ruhaak *et al*[63] and Smith *et al*[64]. One interesting message from these publications is that they all used a different derivatisation method, resulting in identification of different diagnostic product ions which are not compatible with other glycan analysis workflows due

to a variety of reducing-end labels and permethylation/acetylation. The use of a higher-throughput targeted method such as parallel reaction monitoring (PRM[65]) for glycan analysis has not been described in the literature yet but as the field moves to faster scanning mass-spectrometers, more ion glycan targets and their product ions can be monitored with no reduction in instrument performance.

Aims

Overall, we aimed to develop an optimised method suitable for analysis of glycans released from proteins in a complex glycan mixture by LC-MS/MS (summarised in Fig 3.1). This involves the following sub-aims:

1. Evaluate ESI spray voltage for stable negative ion production
2. Optimise the accumulation of ions for suitable glycan quantitation and charge state determination
3. Evaluate acquisition methods for LC-MS glycomics experiments
4. Develop an optimised method suitable for complex released glycan mixture analysis by LC-MS

3.2 Results and Discussion

3.2.1 Optimisation of ESI spray stability

For a glycomics experiment, detection and characterisation of glycan structures is based upon their separation and detection. As ESI is the link between the LC system and the MS, optimisation of this is essential for a successful glycomics experiment. A typical first step in optimising ESI is the direct infusion of tune mix. Using the manufacturer's tune mix and heated electrospray ionisation (HESI) parameters following automatic tune optimisation with the Thermo Tune Plus program, our spray stability was deemed excellent (<5% RSD in TIC intensity), when the tune mix was analysed in the negative ion mode, suggesting the automatically optimised HESI parameters (HESI temperature, spray voltage, sheath gas and auxiliary gas) were suitable for analysis (Table 1)

Heated ESI parameter	**Initially optimised value**
Capillary Temperature (C)	275
Source Heater Temperature (C)	45
Sheath Gas Flow (arbitrary units)	7
Auxiliary Gas Flow (arbitrary units)	0
Sweep Gas Flow (arbitrary units)	0
Source Voltage (V)	2700

Table 3.1 Automatically optimised HESI parameters used as a starting point for manual optimisation

Released glycans from bovine fetuin glycoprotein standard was analysed with our standard PGC-LC-ESI-MS/MS method [29]. When using these previously tuned HESI parameters with the Velos Pro, the eluted glycan peak shape was found to be jagged in a chromatographic run (Fig 3.2). The stability of the signal with an aqueous buffer was then investigated by directly infusing the elution buffer (10mM ammonium bicarbonate in Milli-Q water (Buffer A)) into the HESI source, as the buffer was suspected to be causing the issue, and monitoring a contaminant ion (*m/z* 312.00 $[M-H]^-$) originating from the buffer. Fig. 1 shows that this signal stability test resulted in an average 30% relative standard deviation (RSD) in stability of this contaminant ion which means that for observed ion intensity, we can expect an average of 30% standard error from the true ion intensity. As we were aiming to develop an optimised glycomic analysis method, this error was not suitable for a quantitative application. Furthermore, this value was much higher than expected, when comparing to previously

spray stability measurements of directly infused tune mix which had an average 1.5% RSD in signal stability (Supp. Fig 3.1).

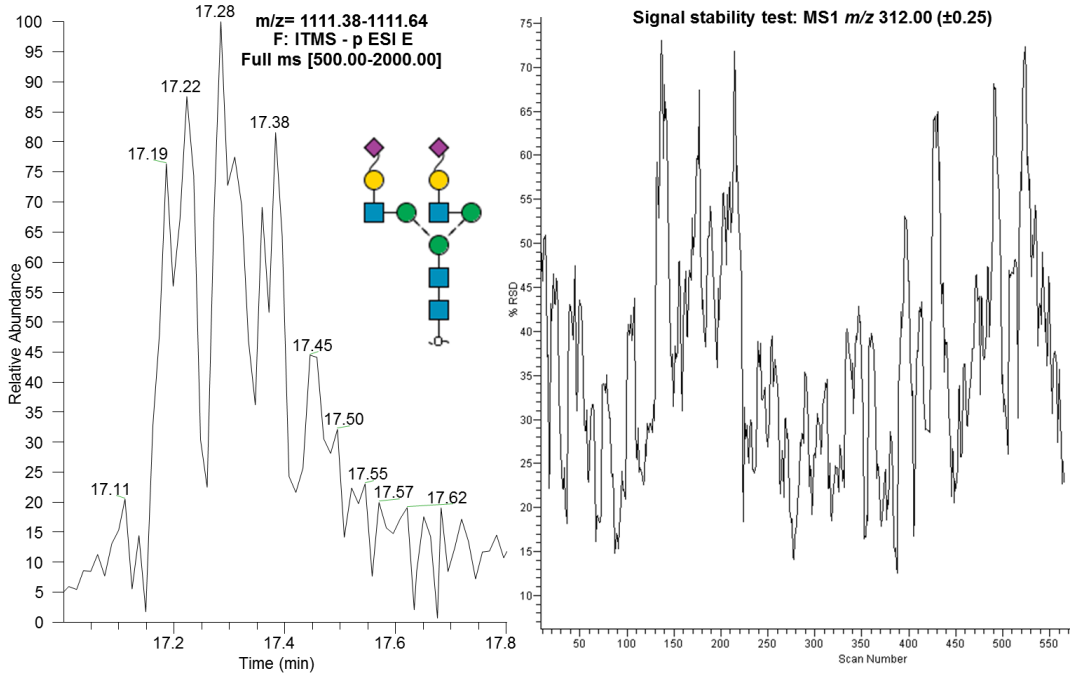


Figure 3.2 Insufficient spray stability with aqueous buffer A (10mM ammonium bicarbonate in Milli-Q water) and default ESI parameters. On the left, representative glycan peak shape of extracted ion chromatogram (*m/z* 1111.38-1111.64) and on the right, spray stability of a contaminant ion (*m/z* 312) monitored by direct infusion (5ul/min) of buffer A used for glycan separation.

One significant difference, expected to be the source of variation in the signal stability test, between the infusion of the standard tune mix and the ammonium bicarbonate buffer was the solvent composition. The solvent composition for the tune mix (50% acetonitrile/25% methanol/24% water/1% acetic acid (v/v)) is 75% organic solvent while buffer A is entirely aqueous (ammonium bicarbonate in 100% Milli-Q water). This alone made it more difficult to obtain a stable spray due to the difficulty of obtaining the constant current indicative of a stable spray[35]. As the chromatographic separation of glycan structures by PGC is dependent on aqueous LC separation buffers (10mM ammonium bicarbonate with increasing acetonitrile composition), we cannot consider adding methanol to improve spray stability without reducing glycan separation.

Several variables that can also influence the stability of the ESI signal include the applied voltage, nebulising gas flow rate, the distance between the spray capillary and ion inlet and the solution flow rate/viscosity/dielectric constant. All of these variables, with the exception of the applied voltage, are fixed due to the available hardware; therefore we focused on the optimisation of the applied voltage for increased spray stability. We found that increasing the spray voltage from -3kV to -5kV resulted in increased spray stability as shown by extracted ion chromatograms and spray stability testing of the previously used (Fig 3.1) aqueous buffer contaminant ion (*m/z* 312.00 $[M-H]^-$) in Fig 3.2. This result reflected the interplay between aqueous composition, applied voltage and constant

current region, demonstrating a correlation of higher applied voltages used with aqueous solvents to obtain increased signal stability[35].

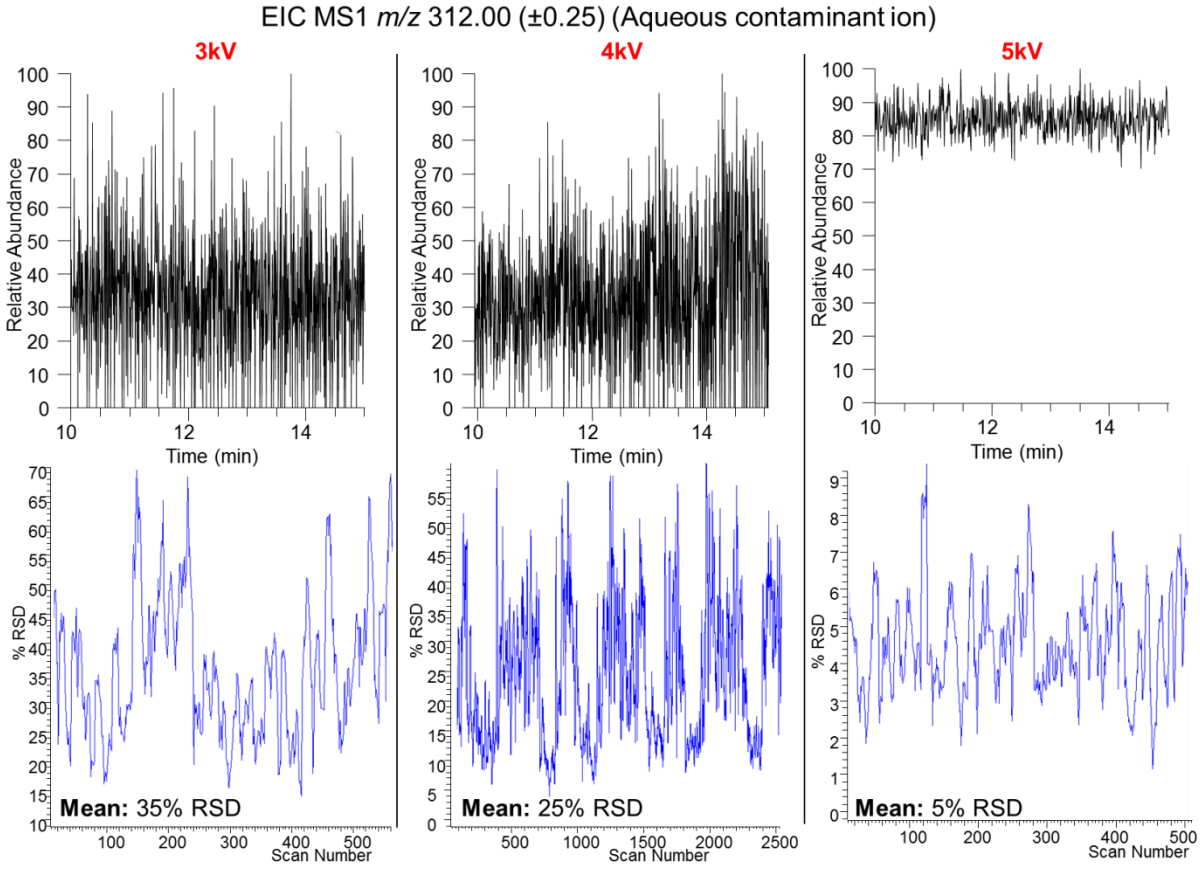


Figure 3.3 Improvement of ion spray stability of aqueous buffer contaminant with increased ESI voltage. Top, extracted ion chromatogram monitoring ion abundance. Bottom, spray stability variation measurements of the buffer A contaminant ion. Both figures from direct infusion (5ul/min) of the buffer A (10mM ammonium bicarbonate in Milli-Q water) used for glycan separation.

While this observation holds true for this contaminant ion, a mix of polysaccharide standards with increasing lengths of mannose, $(Man)_3$ to $(Man)_5$, was used to test the relationship. The same applied voltages with a standard PGC-LC-ESI-MS/MS method[29] resulted in an increase of in-source fragmentation and decreased spray stability and peaks area with increasing spray voltage for the polysaccharide mix used (Fig 3.3). This might be explained by a difference in the properties of the contaminant ion assessed and the polysaccharide standards. Generally, analytes that are highly polar, such as glycans and polysaccharides, have lower electrospray response compared to more hydrophobic analytes due to polar analytes having lower affinity for electrospray droplet surfaces, thereby carrying a lower fraction of the excess charge produced in the electrospray process[35].

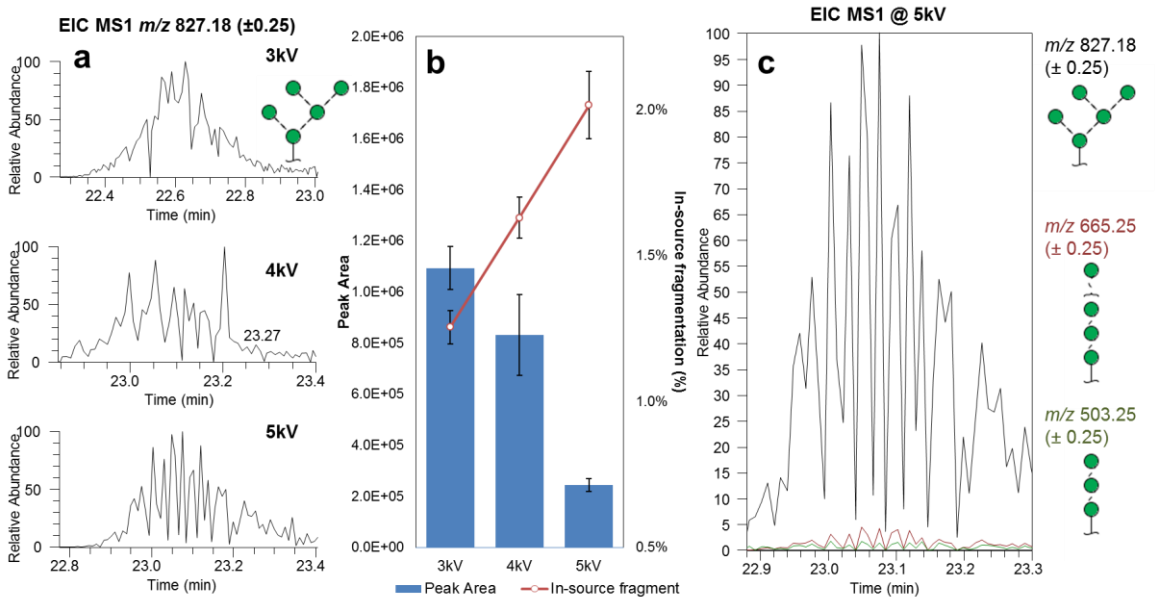


Figure 3.4 Polysaccharide spray stability is decreased with increasing applied spray voltage. a) Extracted ion chromatogram of peak corresponding to $(Man)_5$ with spray voltage from 3kV to 5kV. b) Quantitation of $(Man)_5$ with spray voltage from 3kV to 5kV and the summed peak area of in-source fragments ($(Man)_4$ +$(Man)_3$). c) Enlarged extracted ion chromatogram of $(Man)_5$ at 5kV spray voltage with underlying in-source fragments.

There are two noted reasons for using the lowest voltage possible for negative mode ESI: in-source collision induced dissociation[66] and corona discharge[35]. As reported by Chataigné *et al*, the high voltage of the heated capillary, combined with the labile polysaccharide structure, can result in fragmentation of the glycosidic bonds of polysaccharides[67] and our results are consistent with their observations with the highest peak area of $(Man)_5$ obtained at the lowest applied voltage. The other reason for using the lowest voltage possible is to reduce corona discharge which is a phenomenon which is typically observed in negative ion ESI. This phenomenon results in a higher background and poor ion stability due to electrical discharge from the capillary tip [35]. The use of an electron scavenging gas such as oxygen could also reduce the occurrence of this, allowing higher applied voltages[68], however this was not assessed.

For the above reasons, derivatisation (permethylated hydroxyls or targeted carboxylic acid derivatisation) has often been preferred for glycan analysis as derivatisation of the hydroxyl and sialylated structures results in the carbohydrates no longer possessing a negative charge and thus amenable to positive mode analysis[69]. As permethylation also increases the hydrophobicity of glycans, a greater acetonitrile or methanol composition is required for elution, also causing a more stable spray[70]. Despite these benefits afforded by permethylation for mass spectrometric analysis, such as reduced susceptibility to in-source fragmentation, possible linkage determination and higher ionisation efficiency[71], the PGC-LC-ESI-MS/MS method is used here due to its sample preparation simplicity, resulting in fewer steps and no derivatisation artefacts (including incomplete derivatisation and side-reactions)[29], while still achieving improved structural characterisation as a result of isomer separation. Because of spray voltage optimisation, we found 3kV to be the most

suitable ESI voltage for negative mode carbohydrate ESI and thus use this value for the following experiments, as part of our glycomics method.

3.2.2 Modifying mass-spectrometer parameters to improve data acquisition

While we could not improve peak shape by alteration of ESI parameters, other mass spectrometer-based settings were varied to increase the accuracy of quantitation of the eluted glycans; with one of these settings being the number of microscans set for data acquisition. A microscan (a Thermo-Scientific specific marketing term, also known as trap averaging on Bruker mass-spectrometers) consists of a series of ion pre-injections, pre-scans, ion injections and mass scans, repeated based on the number specified. The instrument then processes these multiple scans and saves the average to the resultant data file as a single scan, thus reducing variation in signal. The use of microscans is not without disadvantage as this results in an increase in instrument cycle time, reducing the number of quantitative points across a peak and/or MS/MS coverage of analytes.

Overall, evaluating the base peak chromatogram of the glycan separation, 2 microscans resulted in smooth enough peaks for automated peak area integration for at least half of *N*-glycans detected in a sample (Fig 3.4). For the lower abundance glycans, the number of microscans had a stronger effect on peak shape, with 3 microscans needed to produce a smooth chromatographic peak with 1 microscan generating a very jagged peak shape. While the use of microscans to address peak shape issues due to unstable signal has not been described in the literature, the importance of the number of microscans has been assessed by Cerdan-Calero *et al*[72] for performance of spectral library-based automatic identification and quantification of metabolites, highlighting the number of microscans as a valuable parameter for reducing scan-to-scan variation. As a result of these optimisations, we settled on 3 microscans in our glycomics workflow.

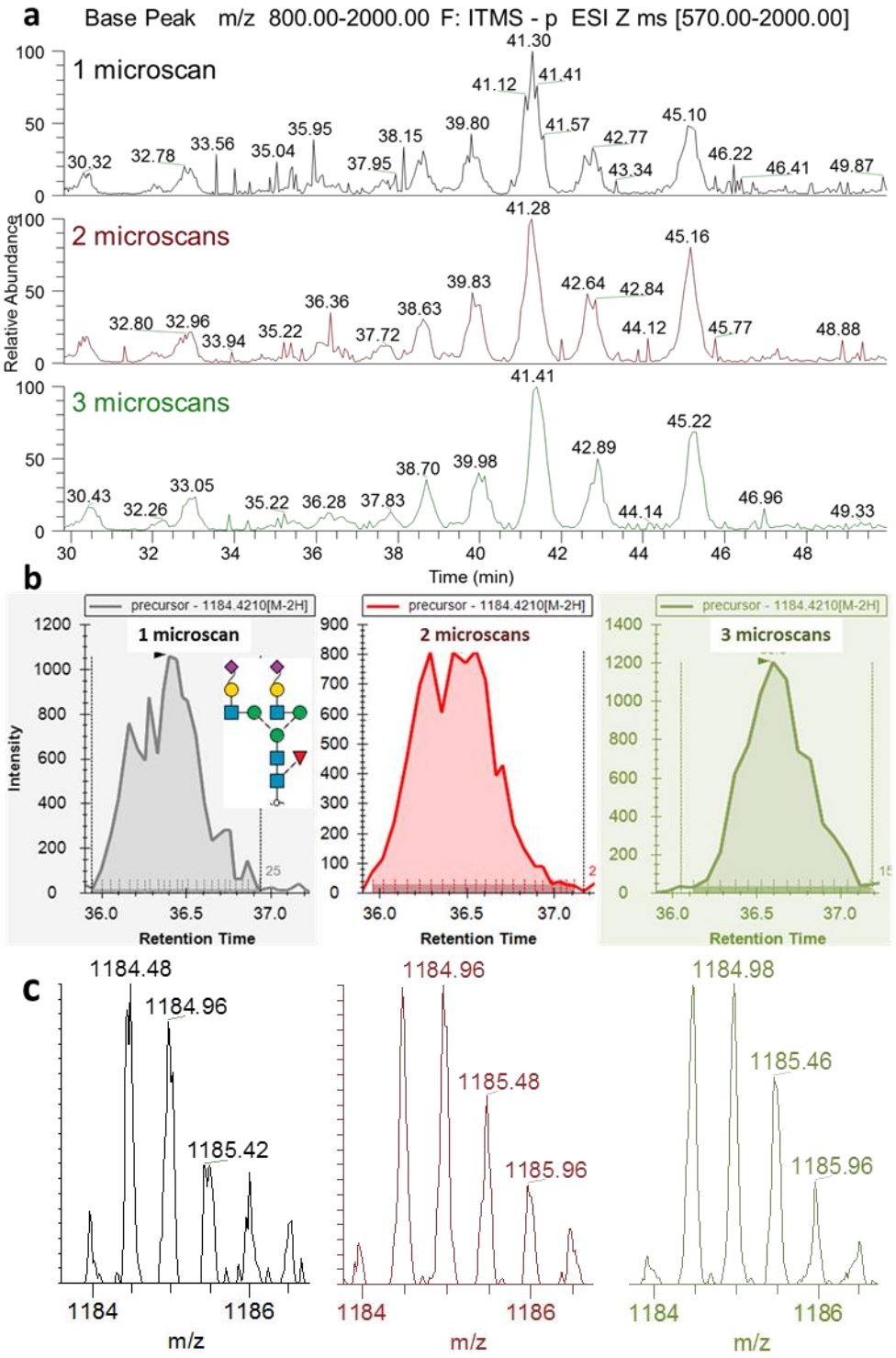


Figure 3.5 Optimisation of glycan MS1 peak shape by modifying MS sampling rate (microscans). a) Base peak chromatogram trace of *N*-glycans released from Human IgG with increasing microscans. b) Effect of microscans on peak shape of less abundant glycans. Dashed lines represent MS1 scans across a peak. c) Isotope distribution of precursor corresponding to each number of microscans.

3.2.3 Optimisation of data acquisition strategy for glycomics

The type of data acquisition strategy and ion accumulation parameters can greatly affect the quality of data acquired and plays a role in the interpretation of this data. The Thermo Velos Pro Ion trap has two types of data dependent MS/MS experiments: double play, in which a full MS1 scan with *m/z*

~0.25 resolution is obtained and full MS2 scans with several *m/z* ~0.35 resolution MS/MS scans follow; and triple play, in which a full MS1 scan with *m/z* ~0.35 resolution is followed by a targeted MS1 scan for the precursor ion with *m/z* ~0.25 resolution and then finished with a full MS2 scan with *m/z* 0.35 resolution. The ion accumulation parameters also are important, as they control the number of ions reaching the mass-analyser and therefore the data quality.

Using repeated equimolar injections of a sample of *N*-glycans released from bovine fetuin, the type of data dependent MS/MS experiments and accumulation time both had an effect on mass spectrum peak shape(Fig 3.5A) and therefore also affected the automatic charge state determination by the mass spectrometer software(Fig 3.5B). As two parameters, accumulation time and number of ions (automatic gain control (AGC)), affect the ion populations, AGC and accumulation time values were increased to the same magnitude. Under-accumulation was observed in the mass spectrum for triple play with 50ms accumulation time, shown by the jagged mass spectrum peaks, which may be responsible for the 35% of unassigned charge states. Over-accumulation was also observed, through space charge effects resulting in reduced mass spectrum resolution, with double play using 100ms accumulation time, demonstrated by the lower resolution mass spectrum peaks, resulting from space-charge effects in the ion-trap[37], but this did not appear to have an effect on the charge state determination. Overall, a double play experiment with 50ms accumulation was observed to be the best balance between accumulation time and subsequent charge state determination.

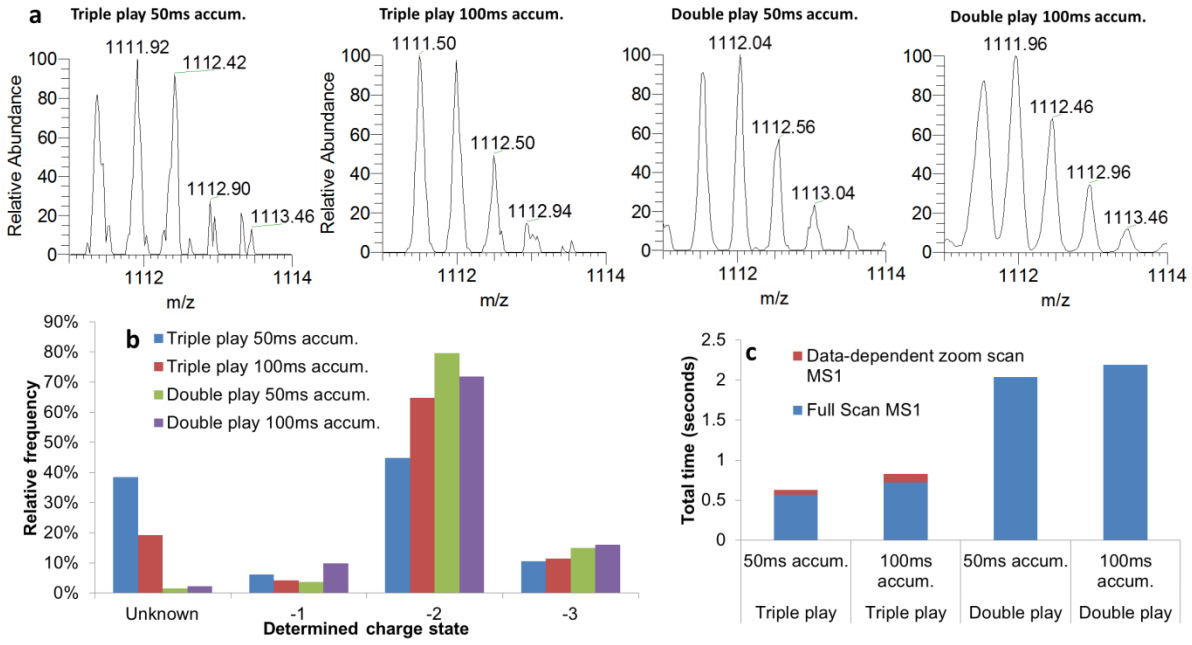


Figure 3.6 Acquisition method optimisation and resulting charge state determination for a single *N*-glycan released from bovine fetuin. a) Mass spectrum peak shape of a single *N*-glycan composition ($(Hex)_5$ $(HexNAc)_4$ $(NeuAc)_2$) with various acquisition methods and ion accumulation. b) Effect of acquisition method and ion accumulation on charge state determination. c) Scan duration depending on acquisition method.

The exact cause of the difference between these types of data-dependent experiments could not be determined beyond the accumulation time as both techniques accumulated the exact same amount

of precursor ions. For most of the precursors with unknown charge states, as assigned by the instrument software, that were collected by the mass spectrometer as part of a triple play experiment, charge state could be manually assigned. This approach would be laborious to do for a biological experiment which can have a complex mixture of glycans with a variety of compositions.

These results were unexpected as quadrupole ion-traps such as this instrument have been widely used, at low resolutions (<0.5 *m/z*), for shotgun proteomics. Literature review revealed that charge state determination is not as important for proteomics data analysis because peptide MS/MS spectra are typically unique and rarely overlap, allowing database searching programs to still be effective with spectra with unassigned charge states[73]. For our data analysis workflow, we first search for glycan compositions using GlycoMod[74] which is heavily reliant on the ability to obtain an accurate singly charged monoisotopic glycan mass. This can be challenging with unassignedprecursor charge states from the mass spectrometer however there are software-based deconvolution tools available which can improve charge state assignment[75].

As large sialylated glycans can be observed with a -4 charge state, it is optimal to have high enough resolution to identify and quantify these glycans while minimising the time to complete a MS1 full scan [7], thereby allowing more time for MS/MS spectra to be collected and to have a sufficient number of MS1 scans across a glycan peak. The data dependent triple play experiment was attractive as it is results in a faster scan time (Fig 3.5C) but charge state determination and accurate precursor mass can be obtained with the double play experiment in a data independent manner. Furthermore, these data-dependent high resolution triple play scans cannot be used for quantitation as they collect the intensity of the precursor in a discontinuous manner unlike the high resolution full scans from double play. Thus we have decided to use a top 5 DDA method using a zoom scan as the first scan event.

3.2.4 Number of precursor ions to be fragmented per duty cycle

One of the final parameters affecting ESI data acquisition that we wished to investigate was the coverage of a complex glycan mixture with a top 5 DDA MS/MS strategy. A mixture of *N*-glycans sourced from a U87MG cell lysate was chosen for this assessment, featuring a heterogeneous mixture of glycan classes (high mannose, paucimannose, hybrid and complex). Using a 60 minute PGC chromatographic separation, a top 5 double play strategy resulted in characteristic fragmentation spectra for all observed *N*-glycan species with sufficient MS1 signal (>10:1 signal:noise ratio) for quantitation (Fig 3.6)[76]. The elution gradient resulted in co-elution of high-mannose structures, and while MS/MS spectra were collected for each of the co-eluting high mannose species (M6 to M9), and allowed differentiation of high mannose *N*-glycan isomers, more MS/MS spectra

could be obtained of these species by gradient optimisation or using a longer gradient. MS/MS spectra with ion counts less than 500 were typically found to be not useful for confirming composition and structure.

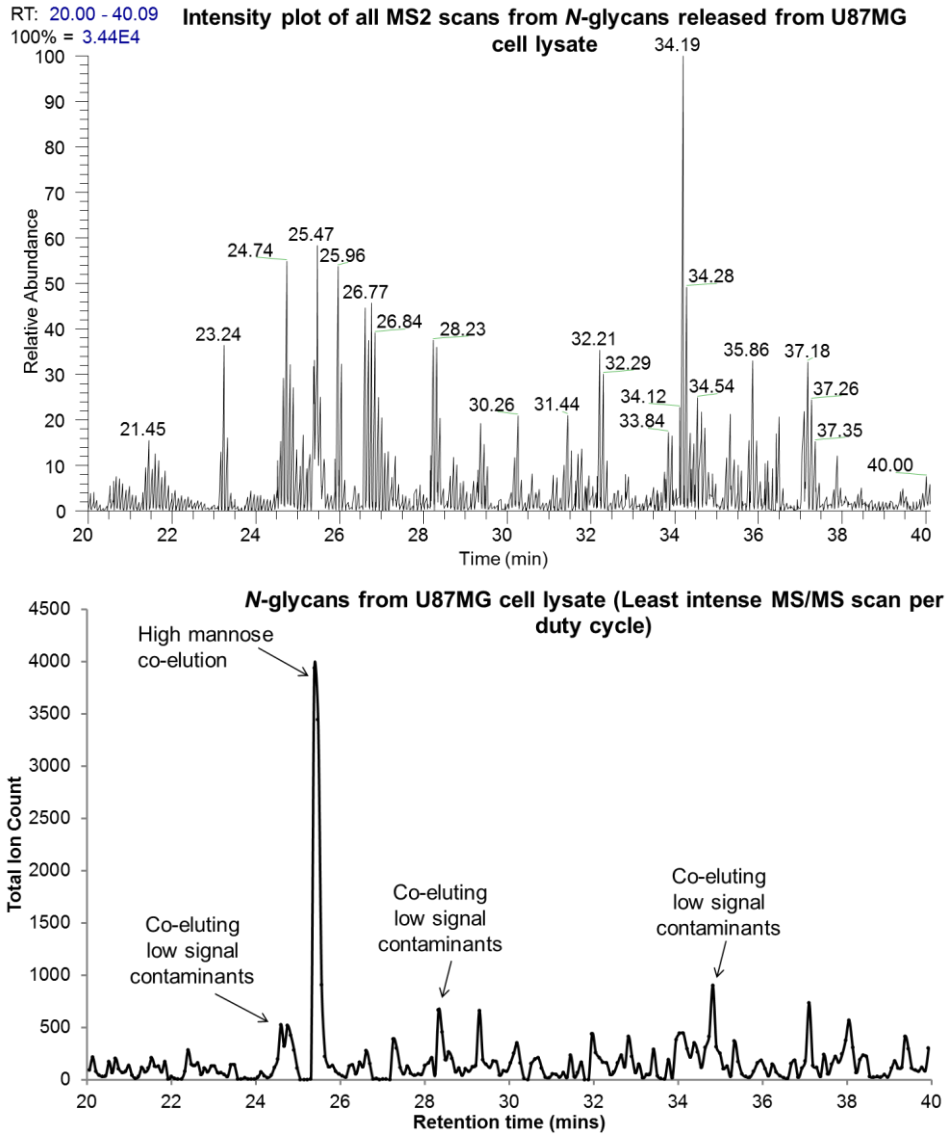


Figure 3.7 Evaluation of top 5 glycan MS/MS coverage of a complex mixture (*N*-glycans released from a human cell lysate). Top, intensity plot of every MS2 scan taken from an analysis of *N*-glycans from a U87MG cell lysate. Bottom, intensity plot of the least intense MS2 scan taken per instrument duty cycle. High mannose co-elution describes multiple co-eluting compositions.

The separation of this glycan mixture was on a PGC column with a 3 micron pore size which produced glycan peak widths of 45 seconds, thereby not requiring a dynamic exclusion policy for glycans. However, if a PGC column with a 5 micron pore size was used, peak widths would be wider which results in lower glycan maximum intensity would be expected to increase the elution overlaps between glycans. A dynamic exclusion policy and/or top6/7 strategy would then be useful in this case and has been implemented in our laboratory by Sakuma *et al* (Unpublished) to be used for PGC columns with a 5 micron pore size.

3.2.5 Development of an optimised duty cycle for glycomics data acquisition

Based on the variety of parameters required for an optimised glycomics MS experiment (ESI, scan resolution, accumulation time and number of MS/MS scans per duty cycle), visualisation of the mass spectrometer duty cycle and scaling of column width to scan time, allowed fine tuning of the data acquisition parameters . Keeping the previous limitations for acquisition in mind (microscans required for low signal variation and double play acquisition method), we have defined a data acquisition method for quantitative glycomics on a linear ion trap instrument that comprises at least ten points of MS1 across a peak and with informative MS/MS fragmentation, by accumulating enough ions and fragmenting them with RE-CID (Fig 3.8).

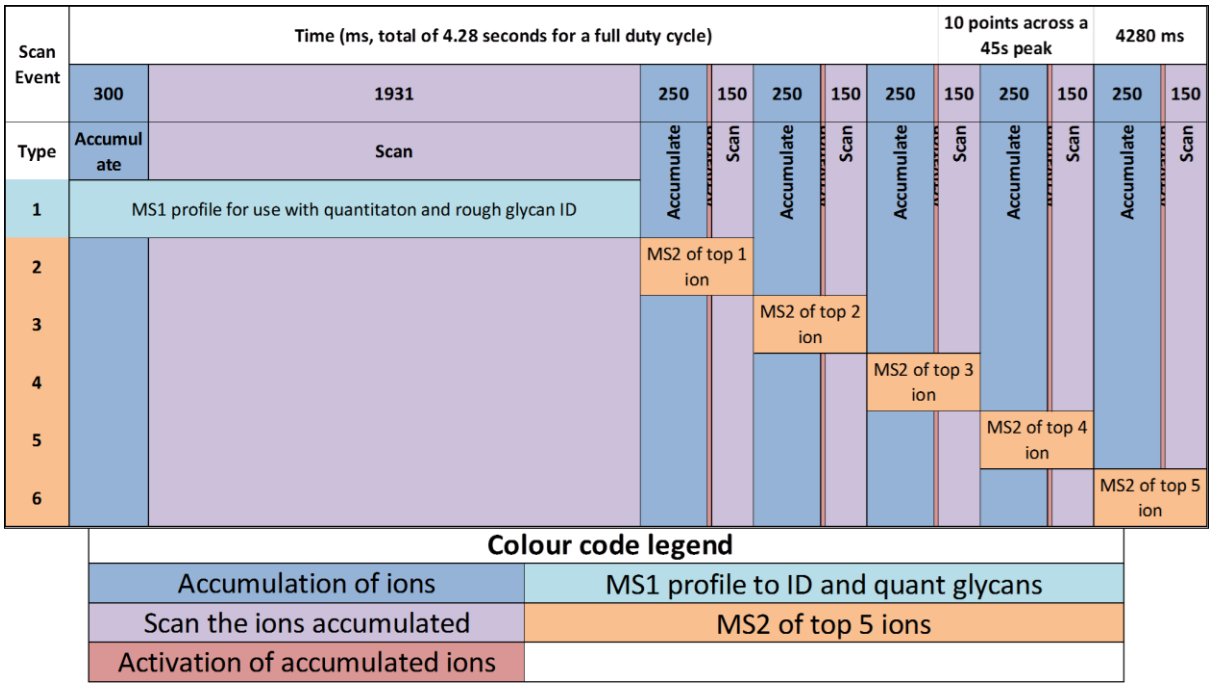


Figure 3.8 Finalised duty cycle based on optimisations performed in this chapter. Table column width is scaled to duty cycle scan time.

This duty cycle estimation, shown in Fig 3.8, was based on values found in the documentation for the Velos Pro mass-spectrometer. If spray stability was not an issue, and 1 microscan was sufficient for smooth chromatographic peak shape, 1400ms would be freed up from the duty cycle (a 30% reduction). While this free time could be allocated towards MS2 or more data points across a peak, a custom, vendor-specific hardware solution would be required to increase spray stability. Using this final duty cycle, with optimised MS parameters from this chapter (Table 3.2), we have established an optimised method for glycan analysis.

Instrument parameter	**Optimised Value**
Spray voltage	2.7 kV
MS1 microscans	3
Type of MS method	Double play with MS1 zoom scan
Resolution of MS1 scan	Zoom (0.25 *m/z* FWHM)
Resolution of MS2 scan	Enhanced (0.35 *m/z* FWHM)
Number of fragmented precursors per duty cycle	5

Table 3.2 Optimised MS parameters for glycan analysis following the sequential optimisation performed in this chapter

3.3 CONCLUSIONS

Despite instrument calibration and tuning with tune mix, vendor recommended ESI parameters were found to be not optimal for glycan analysis with our specific set-up (Dionex LC 4µl/min with a PGC column connected to a HESI source and LTQ Velos Pro). Testing and optimisation of the ESI spray voltage did not improve the negative mode glycan analysis in an aqueous buffer. By modifying the data acquisition parameters to take additional microscans, we improved peak shape (and therefore subsequent quantitation) but with reduced time dedicated for MS2 fragmentation of glycans. The data acquisition method (double play) and ion accumulation values (100ms accumulation time) were also optimised to develop a workflow that has the ability to better determine precursor charge state which was also at the expense of throughput.

Our development of a duty cycle calculator helped optimise many of the parameters of data-dependent LC-MS analysis and to understand the effect of data acquisition strategy choices on duty cycle time. Many of these results are dependent on the intricacies of instrument set-up and are typically not published, but this optimisation paved the way for further technical improvements and biological analysis that are reported in the following chapters. As optimisation was performed using an ion-trap based instrument, these steps can be used to optimise other ion-trap based mass-spectrometers such as Orbitrap, linear ion-trap and 3D ion-trap mass-spectrometers.

Preface

Negative ion CID MS spectra of released glycans can be used for confirmation of glycan composition, sequence and structure. While confirmation of composition and sequence may be relatively straightforward using manual analysis, structural determination is difficult. One key aspect of this difficulty is that diagnostic ions, frequently used for glycan isomer discrimination, are not theoretically structurally-specific by *in-silico* fragmentation methods. Instead, these ions are diagnostic due to their propensity to be produced by one structure over another following fragmentation with defined collision energies at a defined activation time. The optimisation of fragmentation methods and parameters is therefore essential for glycan structural characterisation and slight differences in fragmentation parameters may lead to production of widely differing MS/MS spectra for a single glycan structure.

Following on from the development of a routine method for glycan quantitation (Chapter 3), we present our optimisation of collision energy for glycan fragmentation in our system. We first optimise and evaluate a variety of CID methods available on our ion-trap mass spectrometer. We then compare the fragmentation spectra generated using each optimised method. Variation in MS/MS spectra for a single glycan isomer is then assessed and the applicability of a spectral library approach for glycan isomer discrimination is evaluated. Finally, we compare two optimised CID techniques in their ability to generate diagnostic ions for two glycan structures.

As an example of an unusual glycan modification that was determined by manual interpretation of the released glycan MS/MS fragmentation, a publication is included at the end of this chapter[23]. We utilised PGC-LC-ESI-MS/MS to identify a novel class of *O*-mannose glycans in *Trichoderma reesei* using an alternative CID fragmentation technique which allows detection of product ions with *m/z* values less than 30% of the glycan precursor. We validated the composition, sequence and structure of *O*-glycans found on cellobiohydrolase I (CBHI), an important hydrolytic enzyme for the bioethanol industry. Using alternative characterisation techniques (HPAEC-PAD monosaccharide analysis and C18-LC-ESI-MS/MS glycopeptide analysis), the stereochemistry of these *O*-glycans is determined and glucuronic acid was identified as a terminal residue. These acidic *O*-glycans were found in the catalytic domain of CBHI, with implications for the structure and activity of this enzyme.

4.1 INTRODUCTION

The MS/MS spectrum produced by a fragmented precursor glycan ion is reliant on several key parameters: the glycan ion structure in the gas phase[62], the type of fragmentation method[77] and the parameters defining the fragmentation method. We have reviewed the wide variety of fragmentation methods available for glycomics in the introduction (Paper 1) and thus will only briefly cover them here. There are two main types of fragmentation methods. The first, collision induced dissociation(CID), gives structurally informative product ions resulting from the collision of glycan ions with neutral and nonreactive gases[78]. The second, electron activated dissociation, results from the use of ion-electron or ion-ion reactions to produce glycan fragments typically not observed with collision induced dissociation[79]. CID will be the focus in this chapter, as electron activated dissociation is less widespread in the field of glycomics.

Focusing on collision-induced dissociation, there are a variety of fragmentation methods with each method typically used based on the accompanying mass-analyser. For ion-trap based mass analysers, resonant CID (RE-CID) is frequently used due to its ability to excite and dissociate specifically the precursor mass using resonant RF voltages, resulting in highly efficient precursor dissociation[80]. Time-of-flight and quadrupole -based mass analysers often utilise beam-type CID which is the result of an ion-beam (composed of the precursor ions) being transferred through a collision cell with subsequent product ion detection in the subsequent TOF mass-analyser[81]. One notable limitation of beam-type CID is the inability to perform mass selective activation, resulting in secondary dissociation reactions of the precursor[81]. Instruments such as the Thermo Velos Pro feature multiple CID methods which can generate qualitatively different MS/MS spectra from the same precursor[38]. The ion-trap-based instrument has RE-CID, which is typical for ion-trap-based mass spectrometers, but it also features beam-type CID. This beam-type CID is achieved by passing a precursor ion beam from the ion-trap to the front of the MS at high energy, which results in collision with sheath gas (nitrogen in our case). This ion beam is then transferred back to the ion trap for subsequent analysis[38]. This feature was developed to obtain high efficiency dissociation with no low mass cut-off, which is a limitation with RE-CID[82]. By using a diverse range of CID techniques, complementary MS/MS spectra can be obtained for the same glycan.

While LC-MS/MS experiments generate characteristic MS/MS spectra for each glycan structure present in the sample, the interpretation of these spectra is presently one of the most time consuming steps of analysing glycomics data[83]. Several approaches for automating this process exist but are typically specific for the workflows of the labs in which they were developed such as GRITS, which is limited to permethylated glycan analysis[84], and MultiGlycan, which is limited to positive mode MS glycan analysis[85]. One of the most widely used software packages for *in-silico* fragment prediction is GlycoWorkBench. This software includes prediction of frequently observed

fragment ions generated in typical glycomics experiments (glycosidic bond cleavages B/C/Y/Z and cross-ring cleavages A/X[86]), is compatible with a variety of glycan preparation methods (derivatives, labelled, label-free) and mass spectrometers (high resolution and low resolution mass analysers)[87]. The use of *in-silico* approaches such as this allows annotation of glycan MS/MS spectra, even if a glycan structure has not been observed or described previously. Some isomer discrimination, depending on the presence of theoretical diagnostic ions which are typically linkage-specific cross-ring fragments[86] is also possible.

One limitation of this *in-silico* approach is the absence of application of published diagnostic ions that are product ions specific for particular glycan substructures, as these ions are not highlighted as experimentally validated in the *in-silico* approaches currently available. These diagnostic ions represent specific glycan substructural features and can be produced at a greater abundance compared to the ions produced from glycan structures without these specific substructures, even in some cases from isomeric glycans that produce the same theoretical fragmentation spectra[34]. A collated list of the majority of experimentally determined diagnostic ions for label-free reduced glycan MS/MS analysis has been reported by Everest-Dass *et al* [88]. The difficulty in the application of these ions is the fundamental dependence on competing fragmentation mechanisms that will allow this isomer discrimination, and is therefore reliant on optimised glycan fragmentation[34].

In this chapter, we assess three CID fragmentation techniques available on the ion-trap mass spectrometer and optimise these methods to generate the most informative MS/MS spectra possible for *N*-glycan analysis, including the production of diagnostic ions. In addition, we compare the spectra produced by each method and assess if the fundamentals of spectrum matching, by the generation of a spectral library, is suitable for discriminating the structures of similar glycan isomers. We then demonstrate an application of a variety of CID techniques to characterise a previously undiscovered *O*-glycan which was not possible with the typical glycomics workflow.

Overall, we aim to improve glycan characterisation by CID fragmentation for the purpose of isomer discrimination. This involves the following sub-aims

1. Optimise collision energies for a variety of CID methods
2. Evaluate the similarity in MS/MS spectra produced by each CID method
3. Evaluate a spectral matching approach for glycan isomer discrimination
4. Compare CID methods for generating diagnostic ions

4.2 Results and Discussion

4.2.1 Optimisation of CID fragmentation methods

One of the most important fragmentation settings for informing glycan structure is collision energy, as too low energy can result in insufficient fragmentation for glycan characterisation and too high energy can result in loss of product ions due to ejection of precursor ions which has been found by Webb *et al*[89]. Furthering complicating this, each glycan structure has an optimal energy requirement for dissociation[80]. While instrument vendors have used equations to increase the fragmentation energy with increasing precursor *m/z*, this equation is derived from doubly charged precursor optimisation and frequently cannot be modified on ion-trap mass spectrometers. By empirically deriving the linear collision energy equations for each charge-state, the precursor dissociation energy has been found to be better suited towards the analyte class of interest for increased dissociation efficiency[90].

On the Velos Pro ion-trap we are utilising for this work, three fragmentation modes are available. Resonant CID (RE-CID), Pulsed Q Dissociation (PQD) and trap Higher energy Collision Dissociation (tHCD). While RE-CID and tHCD are widely used and will not be explained further, PQD is a unique fragmentation mode which is a hybrid of both RE-CID and tHCD, where the precursor ions are only briefly resonantly excited. This brief excitation allows product ions to be detected beyond the low mass cut off (LMCO) which is one of the biggest weaknesses of RE-CID, especially as it relates to complete characterization of glycan structure using tandem MS[82].

As shown in Fig 4.1, a core-fucosylated bi-antennary mono-galactosylated *N*-glycan was subjected to varying levels of collision energies with three different fragmentation mechanisms: RE-CID, PQD and tHCD. This *N*-glycan was chosen for optimisation as it features several diagnostic fragment ions. One set represents the core-fucose (*m/z* 350.2 ($Z_{1\alpha}$) and 368.2 ($Y_{1\alpha}$)), which is widely applicable on many instrument platforms due the significant structural difference between core- and outer-arm fucosylation. The other set of diagnostic ions represent the 6-arm antennae extension with galactose (*m/z* 670.3 ($B_4Z_{3\alpha}$) and 688.3 ($B_4Y_{3\alpha}$)), which can sometimes be less applicable due to the structural similarity of arm-extension isomers[91].

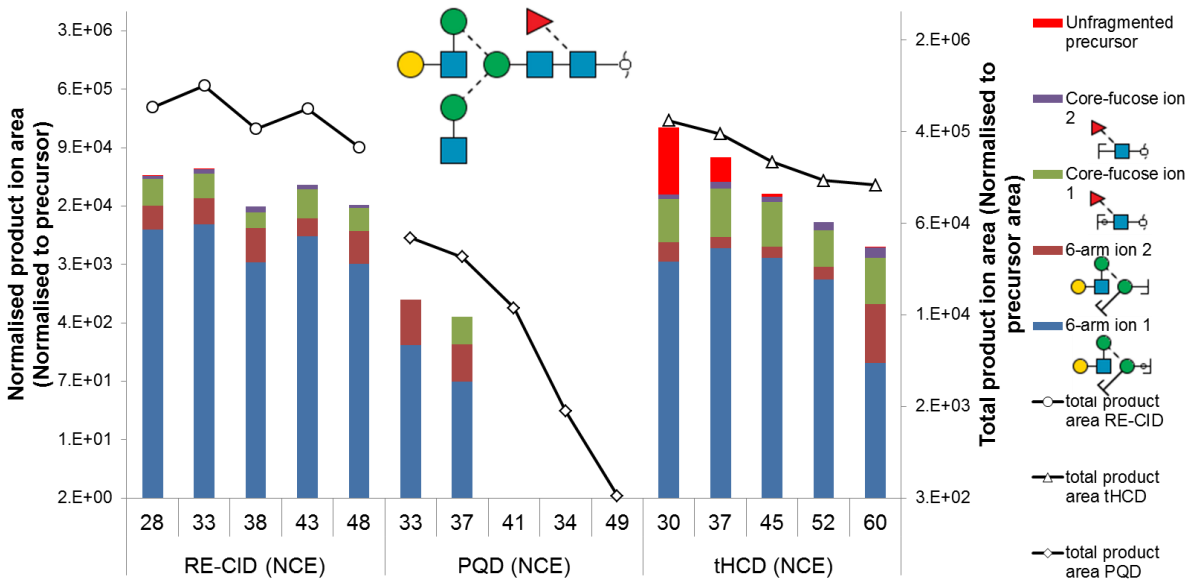


Figure 4.1 Fragmentation optimisation of each technique (RE-CID, PQD and tHCD) performed on a single core fucosylated complex *N*-glycan structure, monitoring diagnostic and total product ion abundance with increasing fragmentation energy. Product ion area based on eight MS/MS scans across a peak.

Comparing total product ion abundances between techniques, overall product ion abundance was generally reduced with increasing collision energy for all fragmentation techniques. As previously mentioned, this was expected due to the possibility of precursor ejection at higher collision energies. PQD had, overall, the lowest product ion intensity for all fragmentation energies, featuring a sharp drop off below 37% NCE, while RE-CID had the most intense total product ion intensity. This was interesting as both RE-CID and PQD feature resonant excitation but PQD has a much shorter excitation to allow the detection of low mass product ions[82, 92]. This shorter excitation likely results in a more intense energy deposition on the precursor and subsequent ejection.

Evaluating the most suitable collision energy for each fragmentation technique, 33% NCE for RE-CID and PQD and 52% NCE for tHCD were observed to be suitable for maximum dissociation of precursor and maximum overall product ion intensity. Looking at specific diagnostic ion abundances across varying collision energies, all diagnostic ions were detected with equivalent ratios between RE-CID and tHCD but PQD was not able to detect core-fucose diagnostic ions (*m/z* 350.20 ($Z_{1\alpha}$) and 368.16 ($Y_{1\alpha}$)) for the majority of collision energies and only one out of the two (37% NCE) useful collision energy levels.

4.2.2 Similarities in MS/MS spectra between CID fragmentation techniques

While optimising the intensity of specific diagnostic ions is important as it allows the discrimination of glycan isomers, overall MS/MS spectra are typically used to confirm glycan structure by mapping specific product ions to the precursor linkage cleavages. Fig 4.2 uses the optimal collision energies derived from Fig 4.1 to evaluate all product ions from seven averaged MS/MS scans, comparing between the three fragmentation techniques. As all experiments were performed on the same glycan structure, same instrument and sequentially in a single batch, we can assess specifically the difference between fragmentation techniques, without needing to account for differences in mass analysers or ion transmission. Fig 4.2A shows that all MS2 look qualitatively similar (especially RE-CID and PQD) but quantitatively, RE-CID and tHCD fragmentation techniques have more intense product ions, as expected from Fig 4.1.

Typical isotopic distributions and mass accuracy values, used to confirm product ion charge state and identity, are demonstrated in Panorama Public (panoramaweb.org/GlycoMS2Diagnostics.url)[93]. We note that glycan structural fragments drawn can be isobaric, i.e. multiple glycan fragments may correspond to a single product ion.

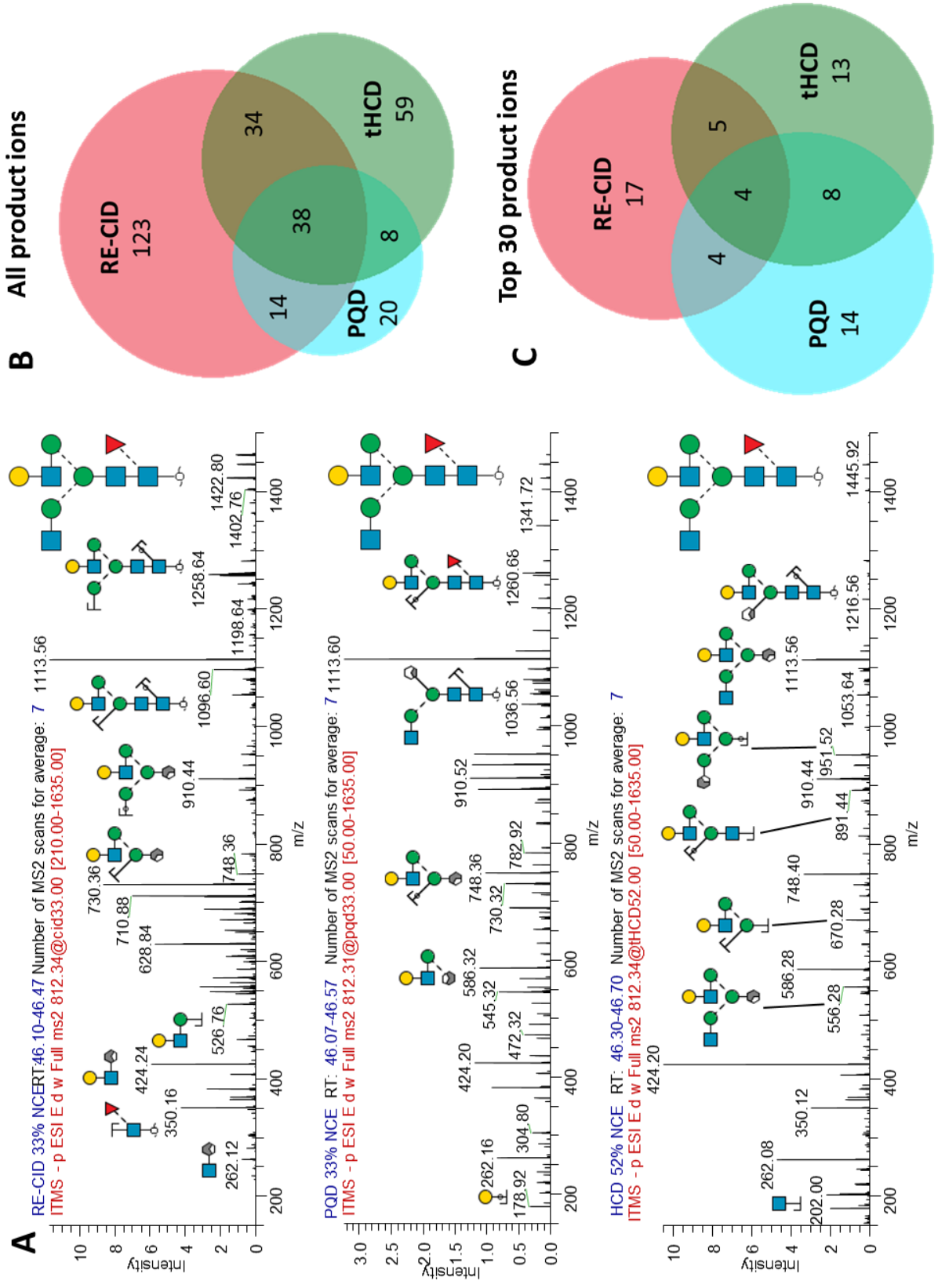


Figure 4.2 Comparison of fragmentation techniques for characterising a neutral core fucosylated complex *N*-glycan structure, using an averaged MS2 spectrum representing 7 individual MS/MS scans. A) MS2 spectra from RE-CID, PQD and tHCD of the same *N*-glycan structure. B) Similarities and differences between fragmentation methods and all their resulting product ions. C) Venn diagram of the 30 most intense product ions from each fragmentation method.

Evaluating each product ion produced by each fragmentation technique, Fig 4.2B shows that RE-CID generates the largest number of product ions, followed by tHCD then PQD. All techniques shared 38 product ions which include many of the abundant ions observed in Fig 4.2A. One contributing reason CID does not detect all product ions observed by other techniques is due to the 30% low mass cut off originating from the mechanism of RE-CID (*m/z* 234 was the lowest product ion observed), whereas PQD (*m/z* 138, with 2 product ions less than 234) and tHCD (*m/z* 98, with 7 ions less than *m/z* 234) have no such low mass cut-off[82]. Of the product ions produced by tHCD and not present in the RE-CID MS/MS spectra, 28 (41%) are the result of multiple bond cleavages and 21 (31%) cannot be assigned by known fragmentation rules or product ion isotopes, highlighting the possibility of gas-phase re-arrangement in the negative polarity analysis with tHCD, which is well known in positive polarity analysis of glycans[48].

The proportion of product ions subject to re-arrangement could in fact be underestimated, in the case of being isomeric with product ions produced through expected fragmentation pathways. If re-arrangement mechanisms are taking place, diagnostic ions would become less specific (structurally diagnostic) as structures could be re-arranged to generate product ions which would not theoretically be produced. This would then result in the production of these ions from glycan isomers without these substructural features, thus reducing diagnostic ion specificity[91].

While we have used all product ions produced as a method of comparing fragmentation techniques, typical MS/MS database searching of biomolecules involves reducing the number of product ions to search, by either taking the most abundant ions, or by applying an ion intensity threshold. We have thus considered only the 30 most intense (top 30) product ions for each technique, to define a limited search space. Overlap between fragmentation techniques is then reduced as shown in Fig 4.2C. PQD and tHCD had a greater overlap (40%) compared to RE-CID. Only 4 product ions (13%) were shared between all techniques. Thus a spectral library approach applied to glycomics, would be quite challenging as different fragmentation mechanisms on the same instrument resulted in widely different top 30 product ion populations. These differences, in addition to other sources of variation between laboratories performing glycan analysis (fragmentation mechanism, collision gas, collision energy and mass-analysers), are a significant reason why the spectral library approach may not be successful for glycomics as compared to proteomics.

4.2.3 Variation of MS/MS spectra and applicability of spectral libraries for isomer discrimination

Further investigating the application of different fragmentation techniques towards the generation of useful spectral libraries, MS/MS spectra produced from a single glycan isomer (F(6)A2[6]G(4)1) were compared to assess spectral variation (Fig 4.3A). Both RE-CID and tHCD have significantly better reproducibility (measured by correlation) between MS2 scans of the same glycan, compared to PQD.

This was in agreement with our collision energy optimisation (Fig 4.1) in which PQD showed greater variation between collision energies compared to RE-CID and tHCD. Assessing correlation for only the top 30 product ions (same as Fig 4.2C) showed no change in this reproducibility for PQD or tHCD however it significantly ($p < 0.05$) improved the reproducibility of CID. This improvement in reproducibility was likely due to the large number of product ions detected for CID (209) compared to tHCD (139) and PQD (80).

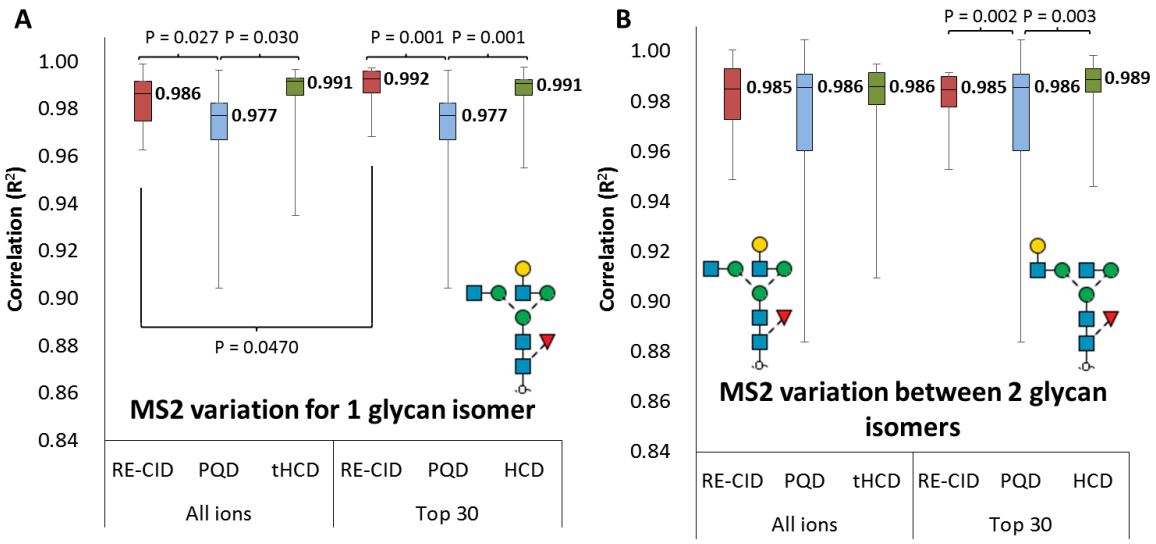


Figure 4.3 Reproducibility of fragmentation techniques between MS2 scans (7 MS/MS scans per glycan isomer). A) Intra-peak comparison for all product ions from RE-CID, PQD and tHCD as well as top 30 product ions B) Inter-peak comparison of MS2 scans between two structurally similar glycan isomers. Bar colours correspond to fragmentation technique (RE-CID (red), PQD (blue) and tHCD (green).

To test our hypothesis of observing lower similarity values (correlation) between MS/MS scans of different glycan isomers due to their differing structure, the same correlation analysis was performed (Fig 4.3B). This involved stepwise comparison between individual MS/MS scans for one glycan isomer and the individual MS/MS scans for the other glycan isomer with the similarity between the MS/MS scans represented as an R^2 correlation value. Correlations performed were also fragmentation method specific.

Comparing MS/MS variation for CID and tHCD, both techniques had slightly reduced reproducibility compared to the correlation within a single peak. For PQD, we found the trend counterintuitive as there was a better median correlation value between MS/MS scans of glycan isomers than within a single glycan isomer. Similar to Fig 4.3A, CID and tHCD had significantly improved reproducibility compared to PQD. This set of glycan isomers were chosen as their RE-CID MS/MS spectra are almost visually identical and quantitatively similar [22] and their structures are typically identified by PGC elution order as well as diagnostic ions. The results from Fig 4.2 and Fig 4.3 demonstrate that a spectral library approach (with no weighting towards specific diagnostic ions), whether comparing all product ions or the top 30 ions, is not suitable for isomer discrimination of these glycan structures. A

later chapter in this thesis will cover one possible solution to this issue, using specific diagnostic ions for correct peak picking instead of the most abundant product ions (Paper 4[22]).

4.2.4 Comparison of fragmentation techniques for generating diagnostic ions

So far for this chapter, we have focused only on one glycan structure, a neutral glycan that is typically found in many mammalian glycoprotein *N*-glycan releases, and thus well studied. Fragmentation can also be used to identify the structure of glycan isomers that are typically not encountered. Fig 4.4 shows a MS/MS spectral comparison between fragmentation techniques for high-mannose glycan structures that originate from a fungal glycoprotein sample primarily composed of cellobiohydrolase I[94]. Fig 4.4A shows two glycan isomers, with complete structure undetermined, appearing as two different peaks due to differential PGC retention. This information alone identifies that there are at least two glycan isomers for this mass and composition. To monitor the abundance of these glycan isomers across multiple sample batches, characteristic product ions from the MS2 of these glycans can serve as criteria to ensure that we are quantifying the same glycan isomer, even when separating these glycans on different stationary phases or different gradients.

We have compared RE-CID and tHCD in their ability to provide characteristic MS/MS spectra for each isomer with unknown, but different structures (Fig 4.4). RE-CID and tHCD were used to produce similar MS/MS spectra (Fig 4.4B/4.4C) with fragmentation energies picked based on results from Figure 4.1 and 4.2 to obtain the most similar MS2 spectra possible. As shown in Fig 4.4B, the first glycan isomer featured an abundant *m/z* 798 product ion (^{2,4}A fragment of the non-reduced GlcNAc in the chitobiose core) and the total MS/MS spectra appeared to be similar between the two fragmentation techniques, aside from the intensity of the unfragmented precursor, resulting from the less efficient dissociation mechanism of tHCD.

For isomer B, *m/z* 900 (Y fragment corresponding to the loss of a terminal mannose) was produced in MS2 spectra from both CID and tHCD. For CID, this product ion was the most intense whereas for tHCD, it was dominated by *m/z* 717 ($C^{0,2}X$ fragment resulting from loss of chitobiose core and cross-ring cleavage of a terminal mannose) and *m/z* 798. The presence of *m/z* 798 in both isomers by tHCD is problematic for discriminating these glycan isomers using the product ions, as the mass is prominent in the MS2 spectra for both isomers. For CID, *m/z* 900 being more dominant in the second isomer compared to the first isomer (17x more intense) and with *m/z* 798.32 being vice-versa (17x more intense), allows these intense product ions to be used for automated discrimination of these glycan isomers based only on the product ions from MS2. Glycan structures were confirmed by characteristic MS/MS fragmentation, with sulfation confirmed using permethylated glycan analysis and high accurate mass analysis (Supp Fig 4.1).

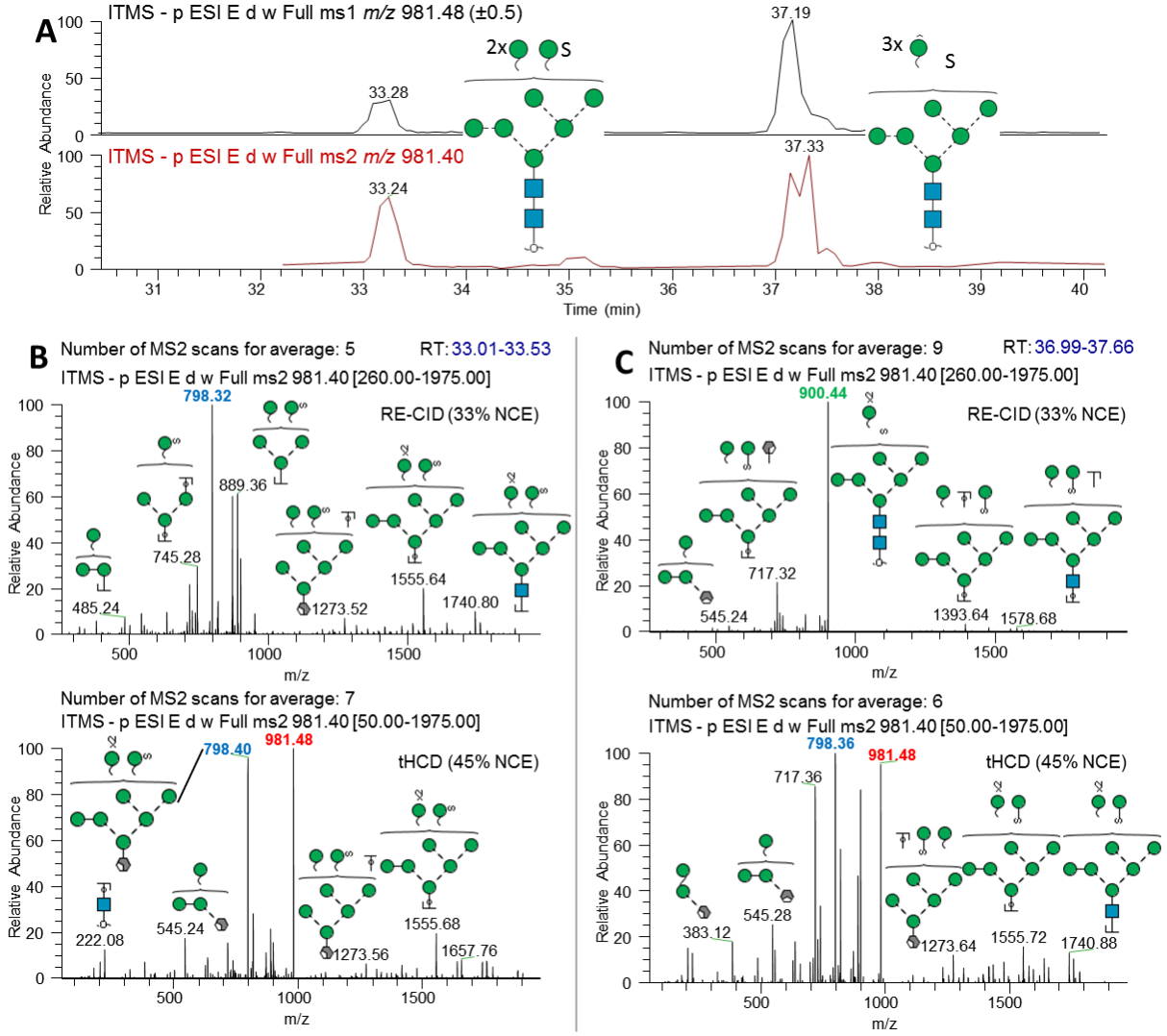


Figure 4.4 Assessing RE-CID and tHCD for generating diagnostic fragments of sulphated oligomannose *N*-glycans. A) Extracted ion chromatogram for both glycan isomers and subsequent MS2 scans. B) Comparison of fragmentation techniques for the first isomer. C) Comparison of fragmentation techniques for the second isomer. Please see Supp Fig 4.1 for confirmation of attached sulphate.

4.3 CONCLUSIONS

A variety of CID techniques, applicable to negative ion trap glycan analysis, were optimised in their ability to fragment a representative *N*-glycan structure efficiently and provide structurally informative product ions. We found RE-CID to fragment the most efficiently, producing the most intense product ions with minimal unfragmented precursor. tHCD performed only slightly poorer and PQD was not able to produce all structurally informative fragments. Each fragmentation technique also produced qualitatively different MS/MS spectra for the same glycan structure and when the 30 most intense product ions were compared between techniques, only 4 product ions were shared by all three fragmentation techniques. This highlights that a top 30 product ion approach, such as that used for proteomic spectral libraries, may require spectral library modification if slightly different fragmentation techniques (or possibly fragmentation parameters) are used.

Evaluating the reproducibility of fragmentation techniques between MS2 scans, RE-CID and tHCD had significantly better reproducibility compared to PQD. Comparing reproducibility between different glycan isomers that are known to fragment differently, the correlation values were not significantly different to those derived from variation within a single glycan isomer. This trend was observed when evaluating all product ions as well as only the top 30 most intense, further confirming a spectral library approach is not suitable alone for differentiating glycan isomers.

We then applied such methods to glycan isomers for which a structure has not yet been assigned. RE-CID produced very different fragmentation spectra for two glycan isomers, producing predominantly *m/z* 798 for the first isomer followed by a dominant *m/z* 900 for the second isomer. tHCD produced a predominant *m/z* 798 for both isomers, highlighting the need to carefully select fragmentation techniques for producing diagnostic ions. Although these diagnostic ions are best suited for discrimination of glycan structure pairs, combinations of diagnostic ions specific for different structural features may overcome this limitation.

4.4 Paper 3 - Enhancing structural characterisation of glucuronidated O-linked glycans using negative mode ion trap higher energy collision-induced dissociation mass spectrometry

An application of a variety of CID techniques to characterise a previously undiscovered *O*-glycan.

Ashwood, C., Abrahams, J.L., Nevalainen, H., Packer, N.H.: Enhancing structural characterisation of glucuronidated O -linked glycans using negative mode ion trap-HCD mass spectrometry. Rapid Commun. Mass Spectrom. 31, 851–858 (2017).

In this paper 3, I performed all experiments, data analysis and writing. Jodie Abrahams assisted in data analysis. Overall responsibility was my own as first author.

Ashwood, C., Abrahams, J. L., Nevalainen, H., & Packer, N. H. (2017). Enhancing structural characterisation of glucuronidated O-linked glycans using negative mode ion trap higher energy collision-induced dissociation mass spectrometry. *Rapid Communications in Mass Spectrometry*, 31(10), 851-858.

DOI: 10.1002/rcm.7851

CHAPTER 5 – AUTOMATING GLYCAN ISOMER DISCRIMINATION USING DIAGNOSTIC IONS

5.0 PREFACE

As described in Chapter 4, negative ion CID MS/MS spectra of released glycans contain ions that are diagnostic for specific structural glycan features and the use of these diagnostic ions can allow discrimination of glycan isomers for the purpose of identification and quantitation. Utilisation of these diagnostic ions for confirmation of glycan structure is completely manual with no guidelines or oversight regarding signal to noise ratio requirements and validation of specificity. Furthermore, differences between vendors and instrument parameters allow the possibility of diagnostic ions to be no longer specific for their originally assigned glycan structure and/or structural feature. This is especially concerning for diagnostic ions which are not theoretically specific for their assigned structure (e.g. Harvey D-ions *m/z* 688 and 670).

In this chapter, we optimise collision energies for a variety of glycan structures based on the abundance of diagnostic ions used for glycan isomer discrimination. Subsequently, we present the use of open-source software (Skyline (https://skyline.ms) and vendor-neutral workflow for identification and utilisation of diagnostic ions for glycan isomer discrimination. Finally, we demonstrate the value of the workflow by discriminating between sialic acid linkage and arm composition isomers using this workflow. To encourage further validation and development of discrimination of glycan isomers using this workflow, we have made the data files and associated Skyline assays available on a freely-available data repository with methods available complying with MIRAGE guidelines.

In paper 4 (Ashwood, C., Lin, C.-H., Thaysen-Andersen, M., Packer, N.H.: Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. J. Am. Soc. Mass Spectrom. (2018)), I performed all experiments and data analysis. Chi-Hung Lin conceived the idea of using Skyline for glycan analysis. Overall responsibility was my own as first author.

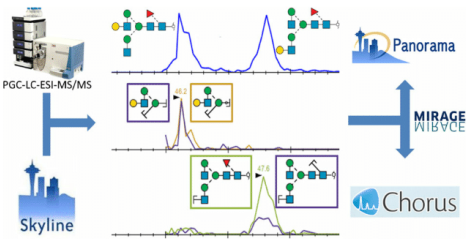
J. Am. Soc. Mass Spectrom. (2018)
DOI: 10.1007/s13361-018-1932-z

FOCUS: MASS SPECTROMETRY IN GLYCOBIOLOGY AND RELATED FIELDS: RESEARCH ARTICLE

Discrimination of Isomers of Released *N*- and *O*-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS

Christopher Ashwood,[1,2] Chi-Hung Lin,[1,3,4] Morten Thaysen-Andersen,[1] Nicolle H. Packer[1,2,4]

[1]Department of Molecular Sciences, Macquarie University, Sydney, Australia
[2]Australian Research Council Centre of Excellence for Nanoscale Biophotonics, Macquarie University, Sydney, Australia
[3]Australian Proteome Analysis Facility, Macquarie University, Sydney, Australia
[4]Institute for Glycomics, Griffith University, Southport, Australia



Abstract. Profiling cellular protein glycosylation is challenging due to the presence of highly similar glycan structures that play diverse roles in cellular physiology. As the anomericity and the exact linkage type of a single glycosidic bond can influence glycan function, there is a demand for improved and automated methods to confirm detailed structural features and to discriminate between structurally similar isomers, overcoming a significant bottleneck in the analysis of data generated by glycomics experiments. We used porous graphitized carbon-LC-ESI-MS/MS to separate and detect released *N*- and *O*-glycan isomers from mammalian model glycoproteins using negative mode resonance activation CID-MS/MS. By interrogating similar fragment spectra from closely related glycan isomers that differ only in arm position and sialyl linkage, product fragment ions for discrimination between these features were discovered. Using the Skyline software, at least two diagnostic fragment ions of high specificity were validated for automated discrimination of sialylation and arm position in *N*-glycan structures, and sialylation in *O*-glycan structures, complementing existing structural diagnostic ions. These diagnostic ions were shown to be useful for isomer discrimination using both linear and 3D ion trap mass spectrometers when analyzing complex glycan mixtures from cell lysates. Skyline was found to serve as a useful tool for automated assessment of glycan isomer discrimination. This platform-independent workflow can potentially be extended to automate the characterization and quantitation of other challenging glycan isomers.

Keywords: Glycomics, Diagnostic ions, Skyline, MS/MS

Received: 7 December 2017/Revised: 30 January 2018/Accepted: 1 February 2018

Introduction

Despite their importance in cell signaling and cell surface biology, glycan structures attached to proteins are frequently avoided as an avenue of investigation, possibly due to their inherent analyte complexity. Glycosylation micro-heterogeneity, the fact that many closely related glycan isomers exist, is an important feature arising from the biosynthesis of glycoproteins and is of proven functional importance, but remains a major analytical challenge in glycomics. The sialyl linkage-type isomers, for example, can influence and define physiology as demonstrated by cancer cells expressing more α2,6- than α2,3-linked sialic acids on glycoproteins, a feature

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13361-018-1932-z) contains supplementary material, which is available to authorized users.

Correspondence to: Nicolle Packer; *e-mail:* nicki.packer@mq.edu.au

Published online: 30 March 2018

that has been associated with increased metastatic potential [1]. Thus, the ability to accurately, rapidly, and sensitively characterize closely related glycan isomers in a quantitative manner is pivotal to further our understanding of glycobiology.

Mass spectrometry is the gold standard to identify and quantitate the relative abundance of structures in mixtures of glycans released from proteins. The monosaccharide composition and the crude topology of glycans can often be deduced solely from tandem MS data. As glycans are generated through an interconnected biosynthetic pathway, many glycans share common substructures such as core-fucosylation and bisecting *N*-acetylglucosamine (GlcNAc). Diagnostic product ions arising from some of these glycan substructures can be used to confirm topology and link the mass spectral data to structure [2]. Combined with powerful isomer separation using porous graphitized carbon (PGC) LC, which affords an orthogonal identifier by providing informative relative or absolute retention times [3], tandem MS analysis is often able to provide sufficient analyte information for a complete or near-complete structural elucidation of isomers of released glycans, but at very low throughput [4].

The applicability of tandem MS analysis for glycan structure determination and isomer discrimination using negative ion fragmentation has, and continues to be, an area of development with Harvey et al. a leader in this area due to their sizeable contributions [5–11]. While these publications cover a wide range of glycan structures with varying properties that affect their fragmentation mechanisms, there are several key notable trends in the mechanism of diagnostic ions being produced. Typical negative ion fragmentation of glycans involves proton abstraction from various hydroxyl groups throughout the glycan [12]. Through charge-remote and charge-induced fragmentation, abstraction of a proton adjacent to a carbon with an adjacent linkage can result in the generation of cross-ring or glycosidic cleavages, depending on the location of the carbon [13].

Evaluating high mannose [6], hybrid, and complex [7] *N*-glycans, Harvey et al. identified product ions that were specific for 6-arm antenna composition through a commonly shared negative ion fragmentation mechanism. The speculated C_{R-1}/Z cleavage, diagnostic for extension of the 6-arm, was produced due to charge-remote fragmentation [6]. Analyzing the same glycan structure with positive ion MS/MS, product ions with the same neutral mass can be observed however these ions were produced by an internal cleavage pathway, thus not diagnostic for 6-arm composition [14]. This example demonstrates that the mechanism of how these product ions are produced is important for generating diagnostic ions for glycan isomer discrimination and a suitable fragmentation method is responsible for the ability to reproduce these fragmentation mechanisms for routine glycan analysis.

One technique to generate informative product ions is the use of resonance excitation collision-induced dissociation (RE-CID), commonly utilized in ion trap mass analyzers. RE-CID is attractive due to extremely high fragmentation efficiency and reproducibility, owing to the control of ion populations by ion trap mass analyzers [15]. Combined with negative polarity (−) mode electrospray (ESI)-tandem mass spectrometry (MS/MS), RE-CID serves as an informative dissociation method for determining monosaccharide constituents and glycan topology through production of both glycosidic bond cleavage products (B-/C- and Y-/Z-ions) and cross-ring cleavage products (A-/X-ions) [16].

Despite past studies identifying RE-CID-MS/MS (−) diagnostic ions for specific glycan substructures [6, 17–19], annotation of glycan fragmentation spectra and subsequent determination of glycan substructures remains largely manual. While several software solutions exist for semi-automated glycan spectral annotation from liquid chromatography(LC)-MS/MS-based experiments [20, 21], these software tools are limited to specific glycan analysis workflows, demonstrating the need for universally applicable software.

We utilize Skyline [22], an open source software platform for LC-mass spectrometry(MS) data analysis, to automatically discriminate between highly similar glycan substructures from a complex mixture of glycans released from both purified glycoproteins and proteins from cell lysates, analyzed using negative mode LC-PGC-ESI-MS/MS. In addition, we provide a freely available, detailed workflow for identifying and characterizing new diagnostic ions for their use in isomer discrimination, providing context for their use in an automated analysis platform.

Methods

Materials

All chemicals were sourced from Sigma-Aldrich (Sydney, Australia) unless otherwise specified. Peptide: *N*-glycosidase F (PNGase F, product #: V4831) was obtained from Promega (Sydney, Australia). All solvents used were LC-MS grade and obtained from Merck Millipore (Sydney Australia). Bovine fetuin (product #: F3385) and human IgG (product #: I4506) were sourced from Sigma-Aldrich (Sydney, Australia). All other chemicals were sequencing grade.

Sample Preparation

The J774A.1 cell line (ATCC® TIB-67) was cultured in a T-75 flask with 10 mL of Dulbecco's modified Eagle's medium supplemented with 10% (*v*/*v*) fetal bovine serum. After the adherent cells reached 70% confluency, the growth medium was removed and adherent cells scraped and collected. The cells were washed with phosphate buffered saline, centrifuged (500 g for 10 min) and then the supernatant was removed, for a total of three washes. The cells were then lysed and protein precipitated using a chloroform/methanol/water extraction (10:10:3, by volume). The precipitated protein was removed, re-solubilized in 4 M urea and protein yield quantified with the Bradford protein assay [23].

N- and *O*-glycans were released from the protein samples as described by Jensen et al. [24]. Briefly, 10 µg of the protein

samples were spotted on polyvinylidene difluoride (PVDF) membranes (Millipore, Sydney, Australia) and stained with Direct Blue (Sigma-Aldrich, Sydney, Australia). The membrane spots were excised and washed in separate wells in a flat bottom polypropylene 96-well plate (Corning Incorporated, NY). *N*-Glycans were released from the membrane-bound protein using 1 U PNGase F (Promega, Sydney, Australia) with overnight incubation at 37°C. Following *N*-glycan removal, 500 mM $NaBH_4$ in a 50-mM KOH solution was added to the membrane spots for 16 h to release reduced *O*-linked glycans by reductive β-elimination.

Released *N*-glycans were reduced with 1 M $NaBH_4$ in a 50-mM KOH solution for 3 h at 50°C, after which the reaction was neutralized by adding equimolar glacial acetic acid. Both *N*-glycans and *O*-glycans were desalted and enriched offline using AG 50W-X8 (Bio-Rad) strong cation exchange followed by PGC solid phase extraction micro-columns (Grace, Columbia, MD, USA) prior to analysis.

PGC-LC-ESI-MS/MS Glycan Analysis

PGC-LC-ESI-MS/MS experiments were performed on an UltiMate3000 high-performance liquid chromatography (HPLC) system (Dionex, Sunnyvale, CA, USA) interfaced with a Linear Trap Quadrupole (LTQ) Velos Pro ion trap (Thermo Scientific, San Jose, CA, USA) unless stated otherwise. Some PGC-LC-ESI-MS/MS experiments were also performed on an 3D ion trap using an Agilent 1100 capillary LC system (Agilent Technologies, Santa Clara, CA) interfaced with an Agilent 6330 LC-MSD 3D Trap XCT ultra.

Separations on both instruments were performed using a PGC-LC column (3 µm, 100 mm × 0.18 mm, Hypercarb, Thermo Scientific) maintained at room temperature and at 50°C for the Agilent and Dionex LC systems, respectively. 10 mM ammonium bicarbonate aqueous solution (solvent A) and 10 mM ammonium bicarbonate aqueous solution with 45% acetonitrile (solvent B) were used as mobile phases. The flow rate was 2 and 4 µL/min for the Agilent and Dionex LC systems, respectively. The same gradient was used for both systems with the following linear gradient program: 0 min, 2% B; linear increase up to 35% B for 53 min; linear increase up to 100% B for 20 min; held constant for 5 min; and then equilibrated at 2% B for 5 min before the next injection—giving a total LC run time of 83 min.

MS Parameters

ESI ion source: the ESI-MS^n analysis was operated in negative ion mode with source voltage at - 3.2 kV for both instruments.

For the Thermo Scientific linear ion trap (LTQ), glycans were analyzed according to these MS conditions: *m/z* 580–2000, 3 microscans, *m/z* 0.25 resolution (FWHM), 5×10^4 automatic gain control (AGC) and 50 ms accumulation time and MS/MS conditions: *m/z* 0.35 resolution (FWHM); 2×10^4 AGC, 300 ms accumulation time, 2 *m/z* window and top five data-dependent acquisition. During MS/MS scans, RE-CID fragmentation was used with helium as the collision cell gas. RE-CID subjected ions to 33% normalized collision energy (NCE) with an activation Q of 0.250 and an activation time of 10 milliseconds (ms). For ion trap HCD (Sup Fig. 7), nitrogen was used as the collision gas. HCD subjected ions to 30–37.5% NCE with a default charge state of 2 and an activation time of 2 ms. Spectral data were acquired in profile mode.

For the Agilent 3D ion trap, glycans were analyzed according to these MS conditions: *m/z* 350–2200, 5 microscans, *m/z* 0.13 resolution (FWHM), 8×10^4 ion current control (ICC) and 200 ms accumulation time and MS/MS conditions: *m/z* 0.13 resolution (FWHM), 8×10^4 ICC, 200 ms accumulation time, 4 *m/z* window and top three data-dependent acquisition. RE-CID acquisition, using helium as the collision cell gas, was performed for characterization of the glycan ions detected in MS/MS scans. Fragmentation amplitude was set to 1 V with "Smart Frag" enabled ramping from 30 to 200% of the fragmentation amplitude for RE-CID with an activation time of 40 ms. All spectral data were acquired in profile mode.

Data Availability

In order to adhere to the MIRAGE guidelines [40], complete LC-MS method parameters are listed as Online Resource 3. The raw data files were uploaded to Chorus (Project 1419). Skyline assays were uploaded to Panorama (https://panoramaweb.org/labkey/project/__r3688/begin.view?) and can be inspected. Raw file names and corresponding Skyline assays are listed in Online Resource 4, Supporting Information.

Data Analysis

Manual Identification of Glycans Lists of ions subjected to RE-CID-MS/MS (−) from relevant spectra of each sample were extracted using the ESI-compass v1.3 Bruker Daltonic Software (Bruker DALTONIK GmbH, Bremen, Germany) and RawMeat v2.1 (Vast Scientific, www.vastscientific.com) for the Agilent 3D ion trap and Thermo LTQ, respectively. After removal of common contaminants, the extracted monoisotopic precursor masses potentially corresponding to glycans were searched against GlycoMod (http://www.expasy.ch/tools/glycomod) to identify putative monosaccharide compositions.

Interpretation and validation of glycan identities were based on the existence of A-, B-, C-, X-, Y-, and Z-product ions consistently found across the majority of MS/MS scans over the elution times of the respective precursor ions. GlycoWorkBench v2.1 (available from https://code.google.com/archive/p/glycoworkbench/) was used for most glycan product ion annotation with an in-house monosaccharide product ion database used to cover fragment gaps. Annotated MS2 spectra for all glycan structures can be found in Online Resource 2.

Skyline Settings for Label-Free Glycomics Skyline (64-bit) v3.6.0.10493 (https://skyline.ms/) was used for all analyses at both the MS and MS/MS level.

Settings Used for MS-Level Experiments – Full Scan Settings For the MS-level filtering, isotope peaks were included by count, only filtering for the monoisotopic peak to be used for precursor ion peak area calculation. The precursor mass analyzer was set to a quadrupole ion trap (QIT) with *m*/*z* 0.5 and *m*/*z* 0.35 resolution (FWHM) for the Agilent 3D ion trap and the Thermo linear ion trap (LTQ), respectively. High selectivity extraction was used and no retention time filtering was performed. MS/MS filtering was used with acquisition method set to "targeted" with the product mass analyzer set to QIT with *m*/*z* 0.5 and *m*/*z* 0.35 resolution (FWHM) for the Agilent 3D ion trap and the Thermo LTQ ion trap, respectively. For Agilent and Thermo QITs, precursor ions were allowed to be in the interval *m*/*z* 50–2200 as the maximum with a method match tolerance of *m*/*z* 0.6.

Mass List Preparation for Skyline The precursor ion targets for Skyline analysis were used from a previously developed list of confirmed glycan composition and included information of (1) the glycan class (paucimannose, high mannose, complex, hybrid), (2) confirmed monosaccharide composition/isomer elution order, (3) experimental precursor ion *m*/*z*, (4) experimental product ion *m*/*z*, (5) precursor ion charge state, and 6) product ion charge state.

Preparation of Data Files for Skyline Import The Agilent ion trap data files were directly converted to .mzml with Proteowizard (version 3.0.10730, available from *proteowizard.sourceforge.net/* [25]) with no filtering applied, then the .mzml file was imported into Skyline(64-bit v3.7.0.10940) [22]. The Thermo LTQ data files were directly imported to Skyline in their .raw format.

MS/MS-Based Peak Picking in Skyline Identification of glycan isomer discriminators was performed by manual analysis of individual MS/MS scans for each glycan isomer in the standards run on the Thermo LTQ. Product ions were deemed to be discriminators by two criteria: (1) the presence in majority (> 80%) of MS/MS scans for the glycan isomer to be discriminated and (2) detection at greater abundance within the spectrum arising from the targeted glycan isomer compared to all other isomers.

Similar to the MS-level peak picking, the Skyline automatic peak picking algorithm was used but only for manual peak picking of glycans not subjected to MS/MS. The integrated peak areas were exported as per MS-level peak picking. Integrated peak areas were exported using a custom report format made for glycan analysis.

Assessment of Specificity In attempts to reduce the influence of precursor ion intensity in the discriminatory specificity calculations of product ions, product ion intensity was normalized to the intensity of the corresponding precursor ion. We use the definition of specificity as the normalized product ion area for the isomer of interest, divided by the normalized product ion area of all isomers observed.

$$\text{Specificity} = \frac{(\text{isomer of interest product ion area}/\text{isomer of interest precursor area})}{(\text{total product ion area for all isomers}/\text{total precursor area})} \quad (1)$$

Skyline was used for area under the curve integration of precursor and product ions. For isomers of interest, a retention window corresponding to the width of each glycan peak (typical width of 45 s) was selected and peak areas corresponding to the precursor and each product ion were calculated and subsequently exported. For total precursor and product ion areas, multiple retention windows were used corresponding to each isomer present in the samples. Using these retention windows, the area under the curve peak area calculations were performed for the isomers of interest, with the additional step of summing the precursor areas (of the most intense charge state) for each isomer as well as the each product ion area.

Results and Discussion

Observation of Potential Product Ions for Glycan Isomer Discrimination

Negative mode ion trap fragmentation of glycans generates multiple fragment types resulting from glycosidic bond cleavages (B-, C-, Y-, and Z-ion fragments) and cross-ring cleavages (A- and X-ion fragments). This broad spectrum of fragments is useful for differentiating glycan isomers [6, 17].

In Fig. 1, we use these diagnostic product ions to confirm that the *N*-glycan is composed of at least one sialic acid residue (*m*/*z* 290.09), the antennae sequence is *N*-acetylneuraminic acid (NeuAc)-galactose (Gal)-GlcNAc-mannose (Man) (*m*/*z* 290.09 (B_1), 470.27 (C_2), 655.29 (B_3), and 817.42 (B_4)) and fucose is attached to the chitobiose core (*m*/*z* 368.23 ($Y_{1\alpha}$) and 553.31 (Z_2)).

In this case, several structural features could not be determined, including on which glycan arm the antenna is located (i.e., α1,3- or α1,6-mannose arm) and the sialyl linkage (i.e., α2,3- or α2,6-). Determination of these specific structural features is more difficult as it requires not only the presence of the correct product ions but also a manual assessment of the relative product ion intensity [17] and, in some cases, the detection of specific cross-ring fragments and their neutral losses [18].

One typical method for demonstrating the ability of diagnostic fragments to differentiate isomers is by interrogating and comparing the MS/MS spectra of two glycan isomers of the same precursor ion mass. We used this method to identify several product ions which could act as diagnostic ions for biantennary *N*-glycan isomers (Fig. 2). As the MS/MS spectra for both glycan isomers was largely similar, subtraction of the

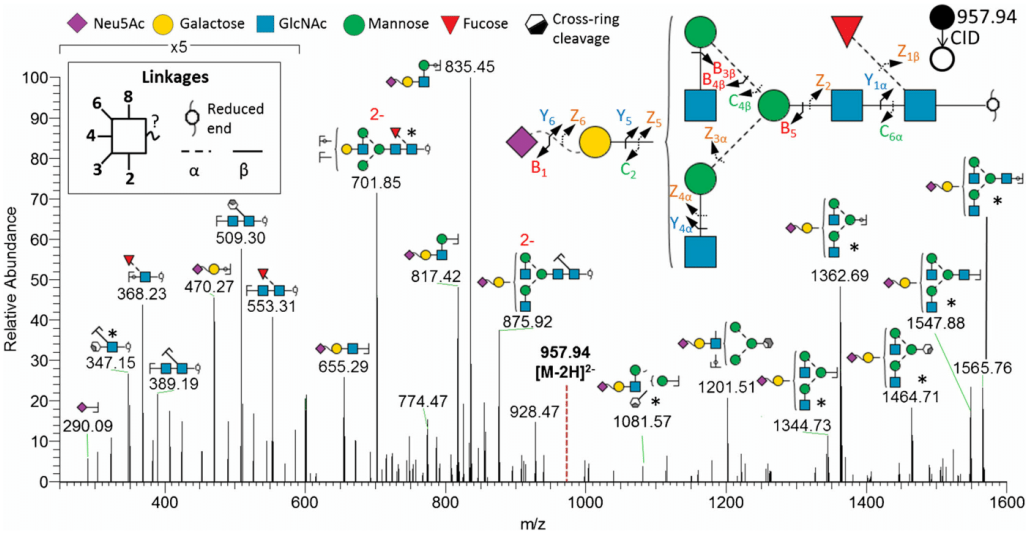


Figure 1. An averaged RE-CID-MS/MS spectrum of a mono-sialylated core-fucosylated bi-antennary complex *N*-glycan. Glycan glycosidic bond cleavages noted in the top right, following Domon and Costello nomenclature. Sialic acid linkage and arm composition could not be determined using the ions acquired. Asterisk (*) denotes structures with more than one possible isomer, only one possible structure per product ion annotated for clarity

spectra was performed to identify ions that are present at greater intensity in one spectrum compared to the other. As a result of this MS/MS spectrum subtraction process of *m/z* 812.37 $[M\text{-}2H]^{2-}$ isomers, we identified *m/z* 304.08 ($B_{2\alpha}{}^{0,4}X_{1\alpha}$), 1260.72($Y_{4\alpha}$), 1281.72 ($Z_{5\alpha}Z_{1\beta}$), and 1445.76 ($Z_{5\alpha}$) as being more abundant in the fragment spectrum for isomer B which possesses a galactosyl extension of the α1,3-linked mannose arm of the bi-antennary *N*-glycan. This structure was also orthogonally confirmed by the PGC elution order [3] with the galactosyl extension of the α1,6-linked mannose arm resulting in an earlier elution compared to an *N*-glycan structure with a galactosyl extension of the α1,3-linked mannose arm. One limitation in the subtraction process is the observation of more intense isotopic peaks for more abundant isomers, which do not correspond to monoisotopic peaks of glycan fragments. To address this limitation, we confirmed these product ions to be monoisotopic peaks through visual examination of the mass spectra and utilization of GlycoWorkBench as an in silico approach to confirm each product ion as a theoretically possible glycan fragment. While this process identified candidate product ions for glycan isomer discrimination, the characterization of product ion abundance and specificity and their relationship with collision energy is important to determine their usefulness as diagnostic ions.

Optimization and Validation of Product Ions that Discriminate Glycan Isomers

As glycans can vary in size, structure, and net charge, which may affect dissociation pathways and therefore the ability to discriminate glycan isomers, optimization of collision energy was performed with several glycan structures of interest (Fig. 3). These glycan structures were chosen as they are typically abundant in *O*- and *N*-linked releases of glycans from mammalian glycoprotein mixtures and cover a wide range of masses (676–2225 Da) and are both asialo- and sialoglycans, which contribute to charge localization and therefore fragmentation pattern [5]. Evaluating first the degree of precursor dissociation, we observed that 28% normalized collision energy (NCE) was the lowest energy required to completely fragment all glycan precursors except for the *O*-glycan (*m/z* 675 $[M\text{-}H]^{1-}$) but the second lowest collision energy assessed, 33% NCE, was effective for complete dissociation of all structures. This could be explained by, following conversion equations to determine the amount of energy applied from NCE% [26], the *m/z* 675.30 ion being subjected to the lowest amount of energy compared to energy used for the heavier ions (Supp. Table 1).

Regarding discriminator abundance, no one collision energy value generated the most intense glycan fragments for all glycan structures. The 48% NCE setting generated on average the lowest amount of product ions. This may be the result of resonant ejection of precursor ions with higher fragmentation energies which has been found by Webb et al. [27].

For the glycan corresponding to *m/z* 675.30 $[M\text{-}H]^{1-}$, discriminator abundance levels were strongly linked to the total product ion intensity, rather than the amount of energy applied. For the glycan corresponding to *m/z* 812.37 $[M\text{-}2H]^{2-}$, varying collision energies changed the relative abundance of several product ions for each discriminator (e.g., discriminator 3 had a relatively low

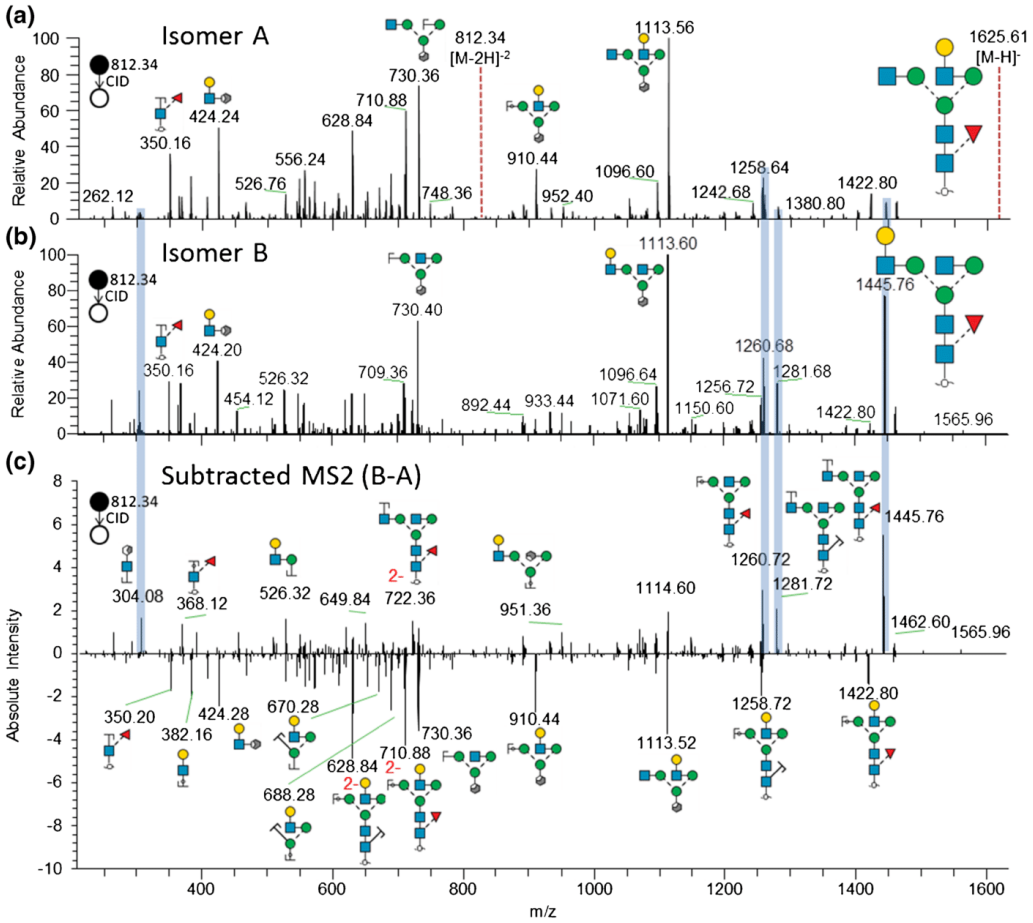


Figure 2. Identification of discriminatory candidate product ions for two neutral *N*-glycan isomers (*m/z* 812.34 $[M\text{-}2H]^{2-}$). (**a**) Average RE-CID-MS/MS spectra for (**a**) isomer A, (**b**) isomer B, and (**c**) isomer B subtracted by isomer A. Only one possible fragment per product ion is annotated for clarity

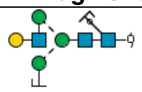
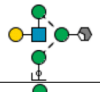
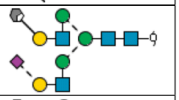
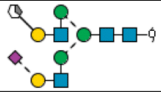
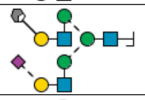
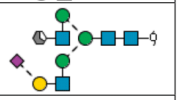
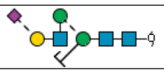
abundance at 38% NCE compared to other assessed collision energies). As a result of this optimization, we chose 33% NCE as the optimal collision energy for complete fragmentation of the precursor ions while producing the most intense product ions.

Discrimination of N-Glycan Arm Extension by Specific Fragment Ions

Previous publications identifying product ions useful for distinguishing glycan isomers often do not feature a quantitative assessment of how specific the chosen product ions are for the isomer that they represent. This hinders the application of these diagnostic ions. We utilize an equation (Eq. 1 above) to give a semi-quantitative assessment of the specificity of the putative discriminatory product ions. In attempts to reduce the influence of precursor intensity on the specificity calculations of product ions, product ion intensity was normalized by the corresponding precursor ion intensity. We use the definition of specificity as the normalized product ion area for the isomer of interest, divided by the normalized product ion area of all isomers observed.

As seen in Fig. 4, we have used published diagnostic ions for core-fucose (*m/z* 350.20 ($Z_{1\alpha}$) and 368.16 ($Y_{1\alpha}$)) and 6-arm composition (*m/z* 670.28 ($B_4Z_{3\alpha}$) and 688.32 ($B_4Y_{3\alpha}$) at 95 and 97% specificity, respectively) to confirm core-fucosylation on both *N*-glycan isomers and identify the glycan structure with 6-arm galactosyl extension, respectively [28, 29]. To identify the other glycan isomer featuring a 3-

Table 1. List of Glycan Fragments for Isomer Discrimination Used for Fragmentation Optimization in Fig. 3 and Their Respective Glycan Structure Fragments (in Cases Where Isomers Exist for a Given Fragment, Only One Structure Is Shown)

Precursor *m/z* (Isomer)	Discriminator #	Product *m/z*	Probable fragment	Precursor *m/z* (Isomer)	Discriminator #	Product *m/z*	Probable fragment
675.30 (B)	1	384.20 $[M-H]^{1-}$		812.37 (A)	1	1258.72 $[M-H]^{1-}$	
	2	631.32 $[M-H]^{1-}$			2	571.32 $[M-H]^{1-}$	
	3	204.08 $[M-H]^{1-}$			3	670.28 $[M-H]^{1-}$	
	4	380.16 $[M-H]^{1-}$			4	910.44 $[M-H]^{1-}$	
	5	222.08 $[M-H]^{1-}$			5	688.32 $[M-H]^{1-}$	
965.96 (A)	1	835.40 $[M-H]^{1-}$		1111.58 (A)	1	1000.6 $[M-2H]^{2-}$	
	2	817.40 $[M-H]^{1-}$			2	1002.00 $[M-2H]^{2-}$	
	3	655.36 $[M-H]^{1-}$			3	889.4 $[M-2H]^{2-}$	
	4	701.84 $[M-2H]^{2-}$			4	898.92 $[M-2H]^{2-}$	
	5	1387.64 $[M-H]^{1-}$			5	891.48 $[M-H]^{1-}$	

arm galactose residue, we have observed glycosidic cleavage product ions specific for this isomer (*m/z* 1445.76 ($Z_{5\alpha}$), 1281.72 ($Z_{5\alpha}Z_{1\beta}$) at 91 and 80% specificity, respectively). The specificities of the product ions for 3-arm position were slightly lower than those routinely used for determination of 6-arm position [29] showing that our shotgun glycomics method is compatible with previously published diagnostic product ions for determining glycan structural features. These previously published diagnostic ions were first detected on a Q-ToF platform using beam-type CID highlighting that this workflow can use diagnostic ions resulting from alternative fragmentation methods and mass analyzers. Through our characterization of the specificity of these diagnostic ions on our platform, we have provided parameters for their use in both manual and automated analysis. These proposed diagnostic ions are of similar intensity and specificity, making them useful in routine analyses.

Interestingly, both structures share core-fucose diagnostic ions but these ions are not represented equally in the MS/MS spectra of both isomers, with the *m/z* 350.20 and *m/z* 368.16 product ions being 70 and 36% specific for the same isomer, respectively. The difference between these two product ions is their type of cleavage with the *m/z* 350.20 ion being a Z_1 cleavage product (cleavage by H^- transfer following deprotonation) whereas the *m/z* 368.16 ion is a Y_1 cleavage product (cleavage by deprotonation) [16]. These results suggest that slight changes in arm composition at the non-reducing end of the glycan affect the production of specific glycan cleavage products at the reducing end which theoretically can provide

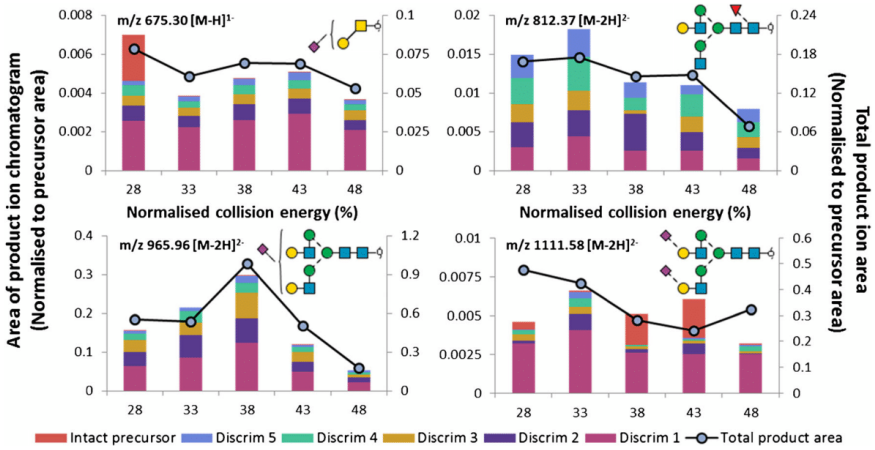


Figure 3. Optimization of CID energy for different *O*- and *N*-glycan structures. Discrimination candidates (detailed in Table 1) and unfragmented precursor ion peak areas assessed as well as total product ion area. All values normalized by precursor peak area. Glycan structures confirmed with MS/MS analysis and sialidase treatment (Supp. Figs. 2 and 3). Raw data can be found in Online Resource 1

candidates for discrimination between isomers with only non-reducing end structural differences.

All product ions assessed were of similar intensity, suggesting these newly identified discriminators for 6-arm position are likely to be observed in any MS/MS spectra where core-fucosylation could be confirmed, therefore not requiring additional accumulation time or increased number of ions for additional structural characterization. As a result of the identification of the 3-arm diagnostic fragments, we can combine these fragments with previously identified 6-arm diagnostic fragments to automatically discriminate and quantitate, by integration of the area under the curve, these glycan isomers in complex mixtures.

Discrimination of Sialyl Linkages by Specific Fragment Ions: N-Glycans

Terminal anionic sialic acids are of great importance in cancer biology. Determination of sialic acid linkages remains an analytical challenge. As shown in Fig. 5, using the candidate MS/MS diagnostic ions for α2,6 sialic acid linkage (*m/z* 835.40 (C_4), 817.40 (B_4), 655.36 (B_3)) and α2,3 sialic acid linkage (*m/z* 979.44 ($B_5Y_{4\alpha}$), 1276.72 ($Y_{6\beta}Y_{4\alpha}$), 1506.80(B_5)), we have identified glycosidic cleavage product ions specific for either α2,6- and α2,3-sialic acid linkages with at least two product ions (*m/z* 817.40/835.40 and *m/z* 979.44/1276.92) having greater than 80% specificity. This specificity value was a goal as it theoretically would be able to correctly discriminate glycan isomers when the glycan isomer of interest is 20% of the peak area of the other isomer, reducing the amount of manual intervention required for integration of glycan isomers.

When evaluating the type of fragments that are specific for each sialic acid linkage, the α2,6-sialoglycan diagnostic fragments are B- and C-type fragments of the antennae whereas the α2,3-linked sialic acid diagnostic fragments are CY-, YY-, and B-type fragments. It can then be hypothesized that changing the sialic acid linkage has an observable effect on the fragmentation pathway that the glycan structure undergoes, thus resulting in differently favored cleavage types that allow discrimination of the two isomeric structures. Charge state of the precursor also can influence the dissociation pathway; however, we have only studied the most abundant charge state as it would be selected more frequently in a data-dependent MS experiment [30].

Previously published sialyl linkage discriminators have been reported to determine the presence of α2,3- or α2,6-linked sialic acid through cross-ring cleavages and neutral losses [18]; however, these ions were either not detected or found to be specifically generated by our analytical platform. As shown in Figs. 1 and 2, the abundance of cross-ring fragments is qualitatively different compared to previously published glycan structure MS/MS spectra [7, 8, 29]. One notable difference is the inability to produce a $^{2,4}A_R$ ion, as described by Harvey et al. notation [11], due to the open-ring reduced GlcNAc making cyclic electron shift unable to occur. In addition to the difference in cross-ring product ions, we have increased intensity of C-type fragments, which have been observed by Harvey to decompose into cross-ring fragments with their beam-type CID platform [6]. We speculate we have reduced the decomposition rate from C-type fragments into cross-ring fragments due to the lower activation energy used in our RE-CID method compared to beam-type CID [3].

Overall, we have found it difficult to pinpoint an exact cause for the slight difference in our MS/MS spectra to

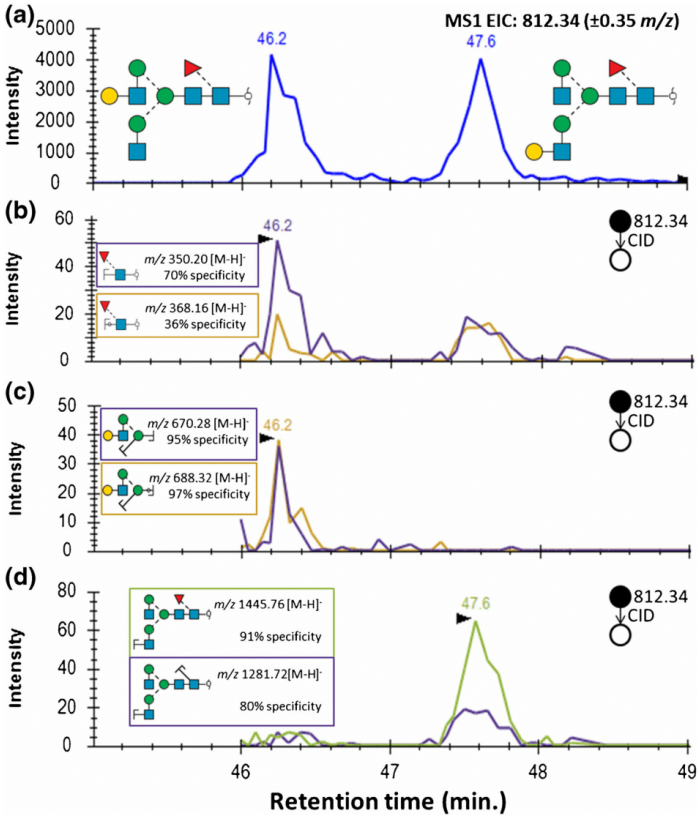


Figure 4. Automated core-fucosylated bi-antennary mono-galactosylated *N*-glycan isomer discrimination using validated and candidate product ions. EICs of (**a**) the glycan precursor (*m/z* 812.34 $[M\text{-}2H]^{2-}$ $(Man)_3(Gal)_1(GlcNAc)_4(Fuc)_1$), (**b**) diagnostic product ions showing core-fucosylation, (**c**) validated diagnostic product ions for 6-arm composition, and (**d**) previously unreported diagnostic ions for 3-arm composition

published spectra by Harvey et al. due to a number of differences between CID methods. For example, between Harvey et al. [29] and the experimental method presented our study, all the MS parameters (except negative polarity) that affect fragmentation differ: collision gas (we used nitrogen, Harvey et al. used argon), CID mechanism (we used resonant excitation, Harvey et al. used beam-type) and activation energy (we used low energy, mV, while Harvey et al. used high energy, V).

By our use of the identified glycosidic bond diagnostic ions generated by RE-CID, we show that we can discriminate sialic acid linkage isomers of *N*-glycans. Similar to the diagnostic product ions we identified for discriminating the 3- and 6-arm mannose antenna extension, the product ions used for discriminating sialic acid linkages are abundant thereby allowing their use for automatic discrimination in a typical shotgun glycomics experiment.

Compared to the mono-sialylated *N*-glycan isomers, bi-sialylated *N*-glycans are more difficult to identify with discriminatory ions because there are four possible sialic acid linkage combinations (α2,3/α2,3-, α2,3/α2,6-, α2,6/α2,3-, and α2,6/α2,6-) as opposed to mono-sialylated *N*-glycans with only two possible linkage combinations (α2,3- or α2,6-). As demonstrated in Fig. 6, using the candidate diagnostic ions for α2,6/α2,6-sialic acid linkage (*m/z* 898.92 ($^{1,5}X_6$)/891.48 ($^{3,5}A_5$) and α2,3/α2,3 sialic acid linkage (*m/z* 1375.76 ($^{2,5}X_6Y_6$)/1295.60 ($^{1,4}A_7Z_3$)), we have identified cross-ring cleavage product ions specific (> 95%) for bi-sialylated *N*-glycans with homogeneous sialic acid linkages. Diagnostic ions with sufficient specificity (> 80%) for the bi-sialylated *N*-glycan containing a combination of α2,3 and α2,6 sialic acid linkages (Fig. 6c) were not able to be identified most likely due to the

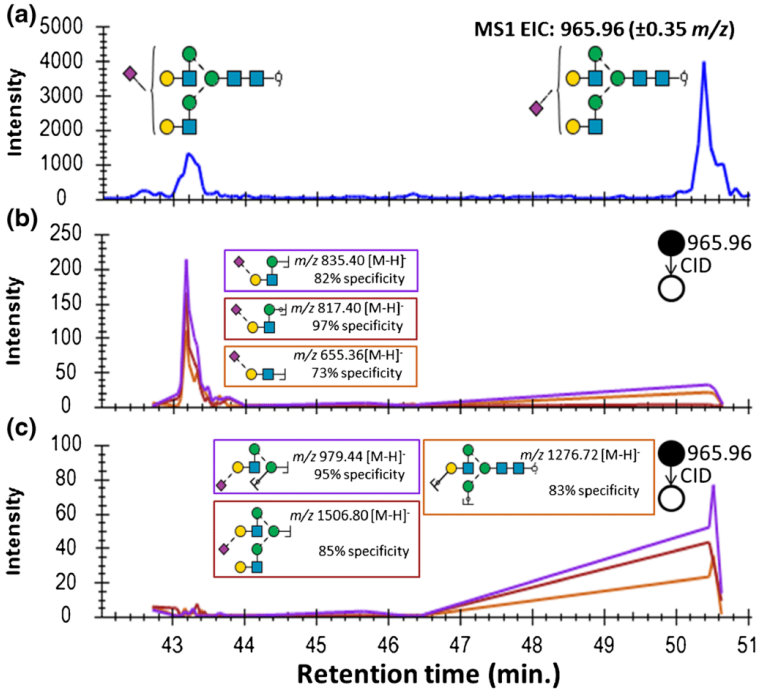


Figure 5. Automated mono-sialylated *N*-glycan sialyl-isomer discrimination using candidate diagnostic product ions. EICs of (**a**) the glycan precursor (*m/z* 965.96 $[M\text{-}2H]^{2-}$ $(Man)_3(Gal)_2(GlcNAc)_4(NeuAc)_1$), (**b**) the previously unreported diagnostic ions for α2,6-sialic acid linkage, and (**c**) the previously unreported diagnostic ions for α2,3-sialic acid linkage isomer

mixture of linkages, with each sialic acid linkage also represented in at least one other isomer.

Unlike the diagnostic ions for the arm position and singly sialylated glycan isomers covered earlier in this manuscript, the most specific fragments for the bi-sialylated bi-antennary *N*-glycans were all the result of multiple cleavages including a cross-ring cleavage. Although these cross-ring cleavage products are not theoretically specific for one glycan isomer, they were still more specific than other fragments generated by glycosidic bond cleavage alone. The ability to produce MS/MS spectra containing product ions resulting from all forms of glycan cleavage demonstrates the versatility of negative mode MS/MS for glycan isomer discrimination.

Discrimination of Sialyl Linkages by Specific Fragment Ions: O-Glycans

O-glycans are a less frequently studied area, despite their importance in breast and colon cancers [31]. Similar to *N*-glycans, *O*-glycans lack a panel of reported and validated diagnostic ions resulting from glycosidic bond cleavage for discrimination of glycan isomers featuring α2,3 sialic acid linkages. As shown in Fig. 7, using the candidate diagnostic ions for α2,6-sialyl linkage (*m/z* 615.32 ($^{1,3}X_{1\alpha}$), 513.24 ($Y_{1\alpha}$) and α2,3-sialic acid linkage (*m/z* 631.32 ($^{1,3}X_2$), 204.08 (Z_1)), we have identified glycosidic and cross-ring cleavage product ions specific (> 95%) for the different mono-sialylated *O*-glycans. Similar to the mono-sialylated bi-antennary *N*-glycan isomers, we were unable to detect previously published [18] discriminatory ions (*m/z* 512.20 ($^{2,4}A_3$), 306.1 ($^{0,4}A_2$-CO_2)) and/or they were at lower abundance and less specific (*m/z* 470.20 (C_2)) than the product ions identified in our study. This further confirms our hypothesis that the method of CID fragmentation can be important since the standard, sialyllactosamine, used for identifying diagnostic ions in the previous study, which used beam-type CID, is structurally similar to the sialylated *O*-glycans.

Assessment of Specificity for Isomer Discrimination

The usefulness of the candidate product ions as diagnostic ions, commonly assessed by comparing MS/MS spectra, does not provide an assessment of their specificity for isomer discrimination on a chromatographic time scale. Our method has its limitations in effectively quantitatively assessing specificity, namely because it was a data-dependent acquisition (DDA) experiment, rather than a selected reaction monitoring (SRM)

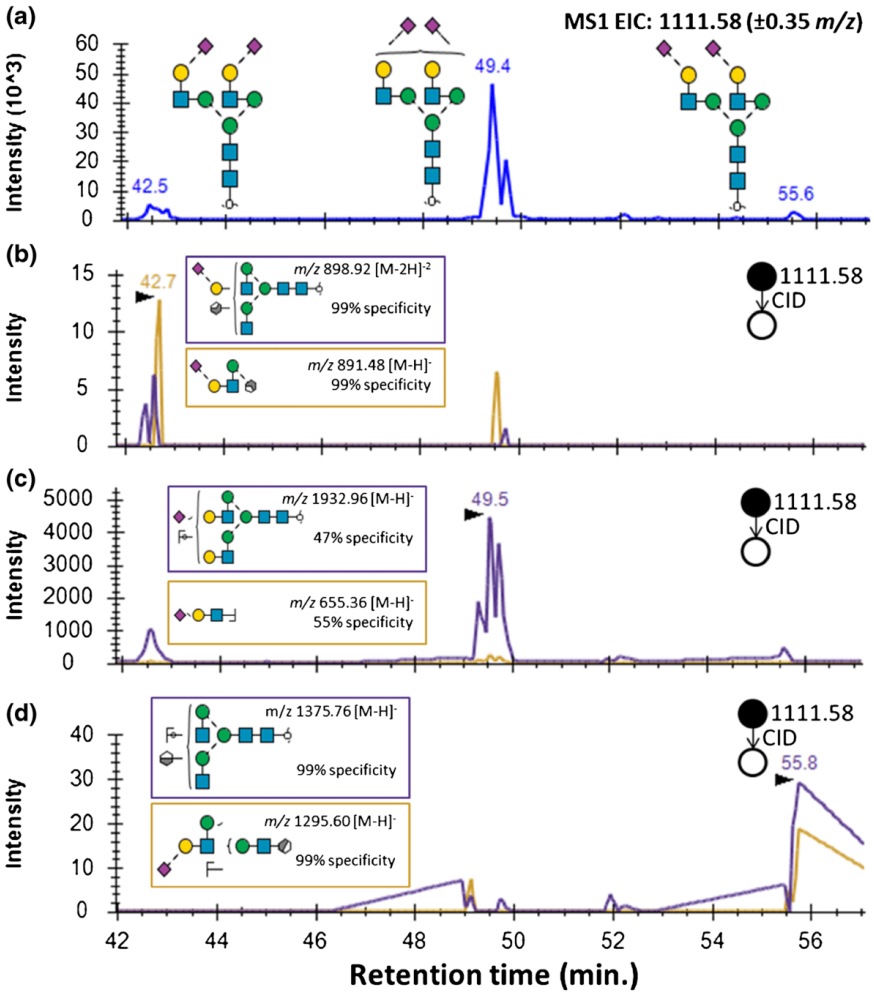


Figure 6. Automated bi-sialylated *N*-glycan sialic acid linkage isomer discrimination using candidate diagnostic product ions. EICs of (**a**) the glycan precursor (*m/z* 1111.58 $[M\text{-}2H]^{2-}$ $(Man)_3(Gal)_2(GlcNAc)_4(NeuAc)_2$). (**b**) The previously unreported diagnostic ions for α2,6-/α2,6-sialic acid linkage isomer, (**c**) the most specific ions for α2,3-/α2,6-sialic acid linkage isomer, and (**d**) the previously unreported diagnostic ions for α2,3/α2,3-sialic acid linkage isomer

or multiple reaction monitoring (MRM) experiment. Despite this, our assessment of specificity is more applicable and relatable to shotgun glycomics experiments.

Our measure of specificity was also very useful for understanding the effect of collision energy on specificity of the generation of candidate diagnostic ions. In Supp. Fig. 1, the specificity of candidate discriminatory ions for 3-arm and 6-arm extension was assessed. With varying collision energies, the specificity of these candidates was found to vary from 70 to 100% with 38% NCE providing, overall, the best specificities for isomer discrimination. Similar to the discriminants for 3-arm and 6-arm position, varying collision energies resulted in specificities ranging from 65 to 100%, and 33% NCE was found to provide, overall, the best specificities for sialic acid linkage isomer discrimination. Similarly to the determination of the arm position (Fig. 3), 48% NCE resulted in the lowest specificity of the identified diagnostic ions, again highlighting the importance of fragmentation energy optimization in

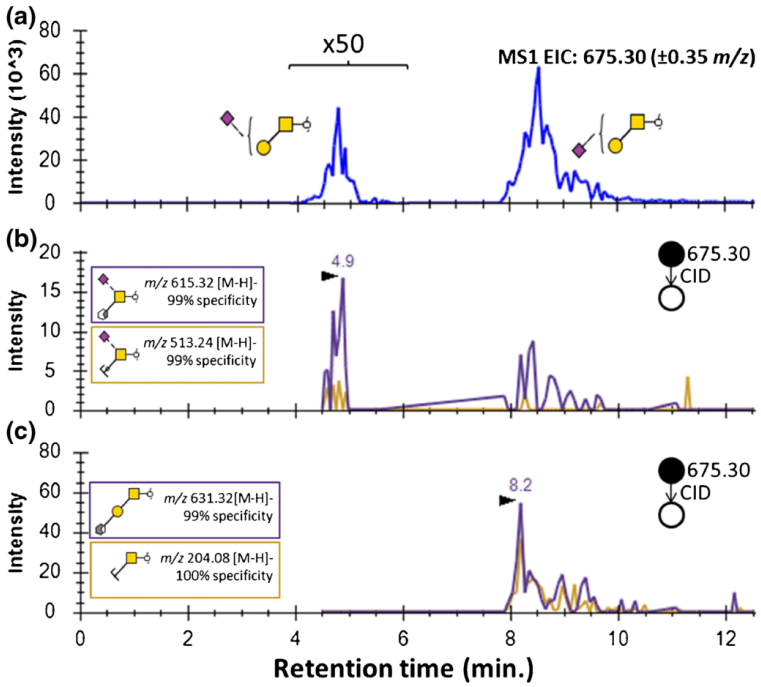


Figure 7. Automated mono-sialylated *O*-glycan sialic acid linkage isomer discrimination using candidate diagnostic product ions. EICs of (**a**) the glycan precursor (*m/z* 675.30 $[M\text{-}H]^{1-}$ $(Gal)_1(GalNAc)_1(NeuAc)_1$). (**b**) The previously unreported diagnostic ions for α2,6-sialic acid linkage isomer and (**c**) the previously unreported diagnostic ions for α2,3-sialic acid linkage isomer

distinguishing glycan isomers. Through this optimization of collision energy, we have determined the best structurally diagnostic ions in a platform able to be used for automated glycan isomer discrimination.

Another finding of interest was for the sialic acid linkage and arm position structural features represented by the diagnostic product ions were not theoretically specific (in silico fragmentation of glycan structure) for their respective isomers (e.g., cross-ring cleavage theoretically specific for one isomer over the other, as utilized by Yu et al. [32]). This is contrasted by the theoretically specific core-fucose diagnostic ion (*m/z* 368.23 ($Y_{1\alpha}$)) used in the structural characterization of the core-fucosylated glycan in Fig. 1. Our hypothesis is that slight differences in the overall glycan structure can result in certain dissociation pathways being favored by one glycan structural isomer over another. Research aiming to understand the fundamental theory of the role of glycan structure on the dissociation pathway has recently become an area of increased focus and could provide more insight into generation of these product ions diagnostic for certain structural features [33].

As demonstrated, we were able to identify diagnostic ions specific for arm position of bi-antennary *N*-glycans and sialic acid linkages of *N*- and *O*-glycans. One key question that remains is the robustness of these diagnostic ions: do these ions remain isomer specific in sample preparation replicates and on other ion trap platforms?

Assessment the Robustness of Diagnostic Fragments

The identification of diagnostic product ions often has not been incorporated into downstream automated software applications, leaving the characterization of glycans with these diagnostic ions to be largely manual and therefore tedious and somewhat subjective. Identification of these diagnostic ions in this study involved post-acquisition manual identification of fragments. For these diagnostic ions to be utilized more widely, fragmentation must be consistent beyond the sample and the instrument platform.

The diagnostic ions for specific arm extension on bi-antennary *N*-glycans were assessed in their use by sample replicates and analyses on other ion trap platforms (Fig. 8). For a sample preparation technical duplicate on the same platform, correct peak picking was successful, demonstrating that these ions are frequently observed and therefore remain useful diagnostics in replicate samples. When that same sample was

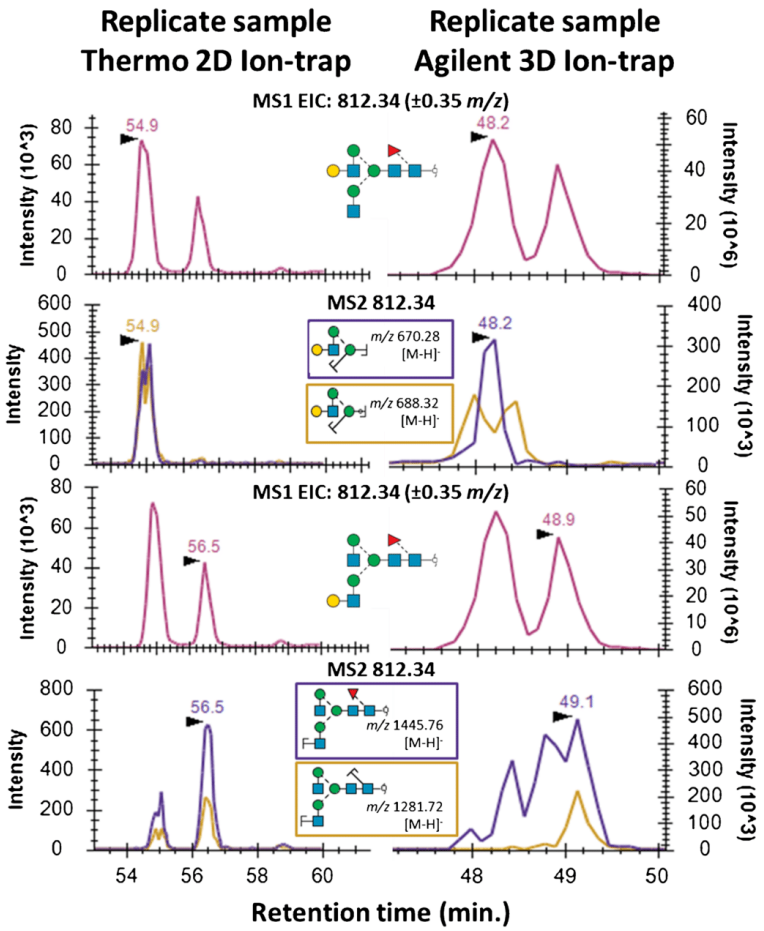


Figure 8. Application of diagnostic product ions for automated peak picking of *N*-glycan 3- and 6-arm composition isomers from technical replicate samples and independent LC-MS platforms. On the left, sample preparation replicate run on the Thermo linear ion trap and on the right, sample preparation replicate run

analyzed on a 3D ion trap, as opposed to the linear ion trap used for identifying these diagnostic product ions, integration of the correct peak was again successful. As shown in Supp. Figs. 4 and 5, this also applies to the other glycans (mono-sialylated *N*- and *O*-glycans) for which diagnostic ions were identified. Mono-sialylated *N*-glycan and *O*-glycan isomers were able to be discriminated using the identified diagnostic ions in both technical replicate samples and on both ion trap platforms.

We found this interesting as the two ion trap platforms feature different instrument architecture (linear vs. 3D) and each platform utilizes vendor-specific parameters for fragmentation (ramping voltage on Bruker and flat NCE for Thermo). Discriminatory ions appeared to be of similar, if not increased, specificity on the 3D ion trap platform demonstrating that these ions are still discriminatory despite optimization of collision energy not being performed on this platform.

As this approach is highly dependent on glycan structure and charge state, small structural changes could result in product ion specificity changes. To investigate if this was the case for sialic acid linkage determination, and see if our methods are applicable to a complex mixture such as a cell lysate, we used diagnostic ions for the mono-sialylated bi-antennary *N*-glycan (Fig. 5) and applied it to core-fucosylated mono-sialylated bi-antennary *N*-glycan isomers (Supp. Fig. 6) released from glycoproteins of a J774 cell lysate. Two of the diagnostic ions previously used were not suitable for these closely related

structures, with one of the ions (*m/z* 835.40, used for discriminating the α2,6 sialic acid linkage) non-specific and the other ion (*m/z* 979.44, used for discriminating the α2,3 sialic acid linkage) not detected in the MS/MS spectra from either isomer. The remaining diagnostic ions (*m/z* 817.40 (B_4), 655.36 (B_3), 1276.72 ($Y_{6\beta}Y_{4\alpha}$), and 1506.80 (B_5)), established for the mono-sialylated *N*-glycan isomers, remained effective for discriminating between the core-fucosylated mono-sialylated *N*-glycan isomers, with surprisingly increased specificity.

This observation could be explained by the fact that the specific diagnostic ions used for discrimination of sialic acid linkages are B- and C-type fragment ions for the α2,6-sialyl linkage isomer and Y-/B- and B-type fragment ions for the α2,3-sialyl linkage isomer; the chitobiose core, where the fucose is attached, is not featured in these cleavage products. The majority of these product ions are however specific for both core-fucosylated and non-core-fucosylated forms of mono-sialylated bi-antennary *N*-glycans, such that identification and validation of candidate discriminatory ions can be performed for each of these glycan isomer types.

For this study, at least two specific product ions per isomer were identified. We have observed that using more than one diagnostic ion for glycan isomer discrimination increases the specificity and thus confidence in characterizing the correct glycan isomer. In the case of more than one structural feature differing between glycan isomers, the use of additional diagnostic ions for each feature could be of benefit.

As automated discrimination of these glycan isomers was performed using instrument platforms set up for discovery-focused shotgun glycomics, this discrimination is compatible with ion trap-based shotgun glycomics experiments, removing the need for targeted pseudo-MS3 analysis ([34]) which has been demonstrated to discriminate similar glycan isomers. To improve the robustness and reproducibility of this workflow for automated glycan isomer analysis, orthogonal data collected from the same run (such as PGC retention time) can also be used to validate the isomer identification [3]. We have also found this method is also compatible with at least one other form of CID, beam-type higher-energy collisional dissociation (HCD). Analyzing the same glycan mixtures with ion trap HCD as the fragmentation method, we have found similar specificities (67–100%) for almost all of the identified product ions (with one exception) to discriminate arm composition and sialylation (Sup Fig. 7).

The Skyline platform was used for assessing specificity and automated peak picking using the identified diagnostic ions, as this software is able to be used with data generated from different vendors and a wide range of mass analyzers. This identifies it as useful software which can be used to implement these fragmentation discriminators for automated glycan isomer discrimination. The use of MS/MS filtering in Skyline with a targeted filtering method was used in this case, although a targeted acquisition method was not used. Unlike previous uses of Skyline for glycomics [35], retention time filtering was not used and was not required for correct peak picking. We see this as a benefit as the use of explicit retention times, in combination with glycan mass, to define glycan structural isomers can lead to erroneous or missed glycan structure quantitation, as column age can result in altered elution characteristics on a porous graphitized carbon stationary phase [36]. The use of an internal retention time standard, such as dextran, could be used for retention time realignment [37] but this has not been accomplished for non-derivatized label-free glycan analysis.

Many of the previous publications identifying diagnostic ions for glycan structure determination have not had the benefit of data repositories or a place for making chromatograms easily accessible. However, from this publication, all data files and Skyline assays used for the automated glycan isomer discrimination and specificity calculations are available on Panorama [38] (https://panoramaweb.org/labkey/project/home/begin.view?) and Chorus [39] (https://chorusproject.org/pages/index.html). This allows improved dissemination and application of these diagnostic ions in the glycomics analytical community.

Conclusion

Sialylation and arm position are biologically important glycan structural features generating closely related isomers that are difficult to confirm by MS/MS alone. We here report a workflow to identify product ions that can identify these structural features and discriminate sialyl linkage on both *N*- and *O*-glycans and arm position on *N*-glycan isomers. These product ions are structure-specific and can also be substructure specific, in the case of the sialic acid linkage on mono-sialylated *N*-glycans with and without core-fucose. While these diagnostic ions have only been identified for native reduced glycans analyzed in negative mode, this workflow has the potential to be used for identification of diagnostic ions for alternative mass spectrometry-based glycan analysis methods such as derivatized glycans or analysis in positive ion mode.

The Skyline software was used to quantify the specificity of these diagnostic ions and implement them into automated glycan isomer integration. These diagnostic ions were robust enough to discriminate between glycan isomers in a complex mixture and were successfully used to interpret data from both linear and 3D ion traps. This workflow may be able to be extended to other glycan isomers that are currently challenging to discriminate based on MS/MS alone.

Acknowledgements

This study was supported by an Australian Postgraduate Award scholarship. We thank M. P. C., Griffith University, for the useful discussions. We also thank the Skyline development team for creating and developing the Skyline software package. This research project was facilitated by access to the Australian Proteomics Analysis Facility (APAF) established under the Australian Government's NCRIS program. C. A. is funded by an Australian Postgraduate Award. M.T.-A. was supported by a fellowship from the Cancer Institute NSW, Australia.

References

1. Lin, S., Kemmner, W., Grigull, S., Schlag, P.M.: Cell surface a2,6-sialylation affects adhesion of breast carcinoma cells. Exp. Cell Res. **110**, 101–110 (2002)
2. Zaia, J.: Mass spectrometry of oligosaccharides. Mass Spectrom. Rev. **23**, 161–227 (2004)
3. Abrahams, J.L., Campbell, M.P., Packer, N.H.: Building a PGC-LC-MS N-glycan retention library and elution mapping resource. Glycoconj. J. 1–15 Advance online publication. https://doi.org/10.1007/s10719-017-9793-4. (2017)
4. Pabst, M., Bondili, J.S., Stadlmann, J., Mach, L., Altmann, F.: Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. Anal. Chem. **79**, 5051–5057 (2007)
5. Harvey, D.J.: Fragmentation of negative ions from carbohydrates: part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates. J. Am. Soc. Mass Spectrom. **16**, 622–630 (2005)
6. Harvey, D.J.: Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose N-linked glycans. J. Am. Soc. Mass Spectrom. **16**, 631–646 (2005)
7. Harvey, D.J.: Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex N-linked glycans. J. Am. Soc. Mass Spectrom. **16**, 647–659 (2005)
8. Harvey, D.J., Jaeken, J., Butler, M., Armitage, A.J., Rudd, P., Dwek, R.A.: Fragmentation of negative ions from N-linked carbohydrates, part 4. Fragmentation of complex glycans lacking substitution on the 6-antenna. J. Mass Spectrom. **45**, 528–535 (2010)
9. Harvey, D.J., Rudd, P.M.: Fragmentation of negative ions from N-linked carbohydrates. Part 5: anionic N-linked glycans. Int. J. Mass Spectrom. **305**, 120–130 (2011)
10. Harvey, D.J., Edgeworth, M., Krishna, B.A., Bonomelli, C., Allman, S.A., Crispin, M., Scrivens, J.H.: Fragmentation of negative ions from N-linked carbohydrates: part 6. Glycans containing one N-acetylglucosamine in the core. Rapid Commun. Mass Spectrom. **28**, 2008–2018 (2014)
11. Harvey, D.J., Abrahams, J.L.: Fragmentation and ion mobility properties of negative ions from N-linked carbohydrates: part 7. Reduced glycans. Rapid Commun. Mass Spectrom. **30**, 627–634 (2016)
12. Pfenninger, A., Karas, M., Finke, B., Stahl, B.: Structural analysis of underivatized neutral human milk oligosaccharides in the negative ion mode by nano-electrospray MSn (part 1: methodology). J. Am. Soc. Mass Spectrom. **13**, 1331–1340 (2002)
13. Doohan, R.A., Hayes, C.A., Harhen, B., Karlsson, N.G.: Negative ion CID fragmentation of o-linked oligosaccharide aldoses-charge induced and charge remote fragmentation. J. Am. Soc. Mass Spectrom. **22**, 1052–1062 (2011)
14. Harvey, D.J., Martin, R.L., Jackson, K.A., Sutton, C.W.: Fragmentation of N-linked glycans with a matrix-assisted laser desorption/ionization ion trap time-of-flight mass spectrometer. Rapid Commun. Mass Spectrom. **18**, 2997–3007 (2004)
15. Bereman, M.S., Canterbury, J.D., Egertson, J.D., Horner, J., Remes, P.M., Schwartz, J.C., Zabrouskov, V., Maccoss, M.J.: Evaluation of front-end higher energy collision-induced dissociation on a bench-top dual pressure linear ion trap mass spectrometer for shotgun proteomics. Anal. Chem. **84**, 1533–1539 (2012)
16. Domon, B., Costello, C.E.: A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj. J. **5**, 397–409 (1988)
17. Everest-Dass, A.V., Abrahams, J.L., Kolarich, D., Packer, N.H., Campbell, M.P.: Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. J. Am. Soc. Mass Spectrom. **24**, 895–906 (2013)
18. Wheeler, S.F., Harvey, D.J.: Negative ion mass spectrometry of sialylated carbohydrates: discrimination of N-acetylneuraminic acid linkages by MALDI-TOF and ESI-TOF mass spectrometry. Anal. Chem. **72**, 5027–5039 (2000)
19. Spina, L., Romeo, D., Impallomeni, G., Garozzo, D., Waidelich, D., Glueckmann, M.: New fragmentation mechanisms in matrix-assisted laser desorption/ionization time-of-flight/time-of-flight tandem mass spectrometry of carbohydrates. Rapid Commun. Mass Spectrom. **18**, 392–398 (2004)
20. Ranzinger, R., Weatherly, D., Arpinar, S., Khan, S., Porterfield, M., Tiemeyer, M., S. W. York: GRITS Toolbox—a software system for the archival, processing and interpretation of glycomics data. In: Annual Meeting of the Society-for-Glycobiology on Glycobiology. p. 1275 (2015)
21. Yu, C.Y., Mayampurath, A., Hu, Y., Zhou, S., Mechref, Y., Tang, H.: Automated annotation and quantification of glycans using liquid chromatography-mass spectrometry. Bioinformatics. **29**, 1706–1707 (2013)
22. MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., MacCoss, M.J.: Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. **26**, 966–968 (2010)
23. Bradford, M.M.: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–254 (1976)
24. Jensen, P.H., Karlsson, N.G., Kolarich, D., Packer, N.H.: Structural analysis of N- and O-glycans released from glycoproteins. Nat. Protoc. **7**, 1299–1310 (2012)
25. Kessner, D., Chambers, M., Burke, R., Agus, D., Mallick, P.: ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics. **24**, 2534–2536 (2008)
26. Nagy, G., Pohl, N.L.B.: Monosaccharide identification as a first step toward de novo carbohydrate sequencing: mass spectrometry strategy for the identification and differentiation of diastereomeric and enantiomeric pentose isomers. Anal. Chem. **87**, 4566–4571 (2015)
27. Webb, I.K., Chen, T.C., Danielson, W.F., Ibrahim, Y.M., Tang, K., Anderson, G.A., Smith, R.D.: Implementation of dipolar resonant excitation for collision induced dissociation with ion mobility/time-of-flight MS. J. Am. Soc. Mass Spectrom. **25**, 563–571 (2014)
28. Everest-Dass, A.V., Jin, D., Thaysen-Andersen, M., Nevalainen, H., Kolarich, D., Packer, N.H.: Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: innate protection against infection by Candida albicans. Glycobiology. **22**, 1465–1479 (2012)
29. Harvey, D.J., Royle, L., Radcliffe, C.M., Rudd, P.M., Dwek, R.A.: Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. Anal. Biochem. **376**, 44–60 (2008)
30. Mcclellan, J.E., Costello, C.E., Connor, P.B.O., Zaia, J.: Influence of charge state on product ion mass spectra and the determination of 4S / 6S sulfation sequence of chondroitin sulfate oligosaccharides. Anal. Chem. **74**, 3760–3771 (2002)
31. Brockhausen, I.: Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. EMBO Rep. **7**, 599–604 (2006)
32. Yu, X., Huang, Y., Lin, C., Costello, C.E.: Energy-dependent electron activated dissociation of metal-adducted permethylated oligosaccharides. Anal. Chem. **84**, 7487–7494 (2012)
33. Bythell, B.J., Abutokaikah, M.T., Wagoner, A.R., Guan, S., Rabus, J.M.: Cationized carbohydrate gas-phase fragmentation chemistry. J. Am. Soc. Mass Spectrom. **28**, 688–703 (2017)
34. Chen, C.-H., Lin, Y.-P., Lin, J.-L., Li, S.-T., Ren, C.-T., Wu, C.-Y., Chen, C.-H.: Rapid identification of terminal sialic acid linkage isomers by pseudo-MS3 mass spectrometry. Isr. J. Chem. **55**, 412–422 (2015)
35. Loziuk, P.L., Hecht, E.S., Muddiman, D.C.: N-linked glycosite profiling and use of Skyline as a platform for characterization and relative quantification of glycans in differentiating xylem of *Populus trichocarpa*. Anal. Bioanal. Chem. **409**, 487–497 (2017)
36. Pabst, M., Altmann, F.: Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides. Anal. Chem. **80**, 7534–7542 (2008)
37. Guile, G.R., Rudd, P.M., Wing, D.R., Prime, S.B., Dwek, R.A.: A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. Anal. Biochem. **240**, 210–226 (1996)
38. Sharma, V., Eckels, J., Taylor, G.K., Shulman, N.J., Stergachis, A.B., Joyner, S.A., Yan, P., Whiteaker, J.R., Halusa, G.N., Schilling, B., Gibson, B.W., Colangelo, C.M., Paulovich, A.G., Carr, S.A., Jaffe, J.D., Maccoss, M.J., Maclean, B.: Panorama: a targeted proteomics knowledge base. J. Proteome Res. **13**, 4205–4210 (2014)
39. Loo, J.A., Matthews, D.E., Yates, J.R.: Focus on bioinformatics, software, and MS-based "Omics," honoring Dr. Michael J. MacCoss, recipient of the 2015 ASMS Biemann medal. J. Am. Soc. Mass Spectrom. **27**, 1715–1718 (2016)

40. Kolarich, D., Rapp, E., Struwe, W.B., Haslam, S.M., Zaia, J., McBride, R., Agravat, S., Campbell, M.P., Kato, M., Ranzinger, R., Kettner, C., York, W.S.: The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data. Mol. Cell. Proteomics. **12**, 991–995 (2013)

CHAPTER 6 – IMPROVING GLYCAN CHARACTERISATION BY NORMALISED RETENTION TIME ON POROUS GRAPHITISED CARBON CHROMATOGRAPHY

6.0 PREFACE

In the previous two chapters, we demonstrated the use of diagnostic ions for glycan isomer discrimination. In cases where there are greater than two glycan isomers, fragmentation spectra is not informative or more robust glycan isomer discrimination is sought, retention time/ elution order can be used to assign glycan structure identity. One challenge in this approach is the lack of normalisation for retention time values, meaning they are not widely applicable across labs, even if they are using identical instrument set-ups.

In this chapter, we present the addition of a dextran ladder as an internal standard for glycomics experiments that can be used to normalise retention time and mass spectrometer signal variation in glycomics experiments. As discussed in the following manuscript introduction, this approach was first applied by Guile et al[95] on fluorescently labelled glycans separated by HILIC. We demonstrate this same approach can be applied to non-derivatised glycans separated by PGC. We identify the elution properties of the ladder with PGC-LC-ESI-MS/MS and, using the internal standard added to released glycan mixtures, we assign system-independent PGC-based retention values to over 200 glycan structures and build a library of the structures with normalised retention time values and their associated MS/MS spectra. Finally, we implement a spectral library for automated peak assignment of the dextran ladder to allow automated normalisation of variation in separation of glycan mixtures.

6.1 Paper 5(In preparation) – Dextran ladder as an internal standard for PGC-LC-ESI-MS/MS glycomics

Title: Utilisation of a dextran ladder to standardise PGC-LC-MS-based glycomics

Authorship: Christopher Ashwood[1,2], Brian Pratt[3], Brendan MacLean[3], Nicolle H. Packer[1,2]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia
2. ARC Centre of Excellence for Nanoscale Biophotonics, Macquarie University, Sydney, NSW, Australia
3. Department of Genome Sciences, University of Washington, Seattle, WA

Target journal for submission: Analytical chemistry

Abstract

Porous graphitised carbon (PGC) based chromatographic separation of glycans achieves high-resolution separation of glycan mixtures released from glycoproteins, including structurally similar isomers. While there is some understanding of glycan separation on PGC, system-independent retention values have not been established.

Using hydrolysed dextran as an internal standard and Skyline software for post-acquisition normalisation, retention time and glycan peak area variation for replicate injections of glycan mixtures was significantly reduced. Normalisation of retention time to the dextran ladder allowed assignment of system-independent retention values, values that are applicable to all PGC-based separations regardless of chromatographic system. We have thus built a library of over 300 PGC-separated glycan structures with normalised retention values.

To further define the mechanism of glycan separation with PGC, we identified predictive models for the chromatographic effects resulting from addition and/or removal of core-fucosylation and bisecting GlcNAc based on the PGC normalised retention time library. A dextran ladder spectral library was built to ensure correct retention time assignment of the internal standard added to glycan mixtures. Using the spectral matching feature in Skyline, isomeric discrimination between *O*-mannosylated glycans and the glucose-based dextran ladder was achieved.

As a result of these efforts, system-independent assignment of glycan structure based only on precursor mass and glucose unit value can be achieved which was previously not possible using PGC-based LC-MS.

Introduction

Glycomics aims to comprehensively study the entire complement of sugars found in an organism and ideally needs to detect and quantify each individual glycan structure present[96]. The ability to discriminate between the often encountered isomeric glycan structures is necessary to accomplish this aim.

Techniques such as capillary electrophoresis and liquid chromatography (LC) are able to separate isomeric glycan structures[14]. For glycan analysis, hydrophilic interaction chromatography (HILIC) can obtain robust separations[28] whereas PGC is a alternative stationary phase capable of high resolving power for separating glycan isomers[29]. While capable of resolving isomeric glycan structures, the exact separation mechanism of PGC is unknown therefore glycan structure retention times cannot be predicted using in-silico models[97].

To understand the applicability of PGC for separation of glycans, glycan retention libraries and elution order rules have been established[41, 98–100]. As a result of these efforts, glycan structural identity can be assigned to peaks detected in a chromatographic run but the retention times accompanying these glycan libraries are not widely applicable as they are system specific values[41, 98–100]. These variations in observed glycan retention time impede precise structural determination using PGC. In an alternative analytical approach, Guile *et al*[95] used a dextran ladder as a glycan retention index for fluorescent detection of glycans separated by HILIC and this has been further developed over the past two decades into a robust and routine method[28, 101–103].

Through the use of a dextran ladder with an increasing degree of polymerisation, system-independent retention constants for glycans can be derived which are independent of the chromatographic system used, though specific for the column stationary phase. These system-independent retention constants are measured in glucose units (GU) as the dextran ladder is a glucose polymer linked through α1,6 glycosidic bonds[104]. One successful example of its implementation is the construction of GlycoBase, a database of over 375 2-aminobenzamide labelled glycan structures separated on HILIC and assigned with corresponding GU values[105]. This database allows automated structural assignment to observed peaks using fluorescent detection and their GU values but requires reductive amination-based glycan labelling[103].

An equivalent standardisation of retention time on which to build a retention index database does not exist for PGC despite the high resolving power. In addition, the capacity to interface PGC separation of underivatised glycans with mass-spectrometric methods for characterising and quantifying glycans can provide additional value to a glycomics experiment compared to fluorescence-based glycan detection.

In this report, we characterise the elution behaviour of a reduced dextran ladder separated by PGC-LC and use this ladder an internal standard for retention time and peak area normalisation of glycans separated and analysed by negative mode PGC-LC-ESI-MS/MS. A system-independent glycan retention value library was created and subsequently applied to characterise all released *N*-glycans detected from a U87MG cell lysate. To improve the automation and robustness of analysis, a spectral library was created of the dextran ladder to allow high throughput retention time normalisation of complex glycan mixtures even those containing glycans isomeric to the ladder.

Experimental section

Chemicals and materials

All chemicals were sourced from Sigma Aldrich (Sydney, Australia) unless otherwise specified. Peptide:*N*-glycosidase F (PNGase F, product#:V4831) was obtained from Promega (Sydney, Australia). All solvents used were LC-MS grade and obtained from Merck Millipore (Sydney Australia). Bovine ribonuclease B (product#:R7884), porcine gastric mucin (product#:M1778), Human IgA (product#:I1010), bovine lactoferrin (product#:L9507), Human lactoferrin (product#:L0520), bovine fetuin (product#:F3385) and human IgG (product#:I4506) were sourced from Sigma Aldrich (Sydney, Australia). Human neutrophil elastase (product#:342-40) was sourced from LeeBio (Maryland Heights, USA). Fungal cellobiohydrolase I was isolated as previously described[23]. All other chemicals were analytical grade.

Cell line culturing

The U87MG cell line was cultured in a T-75 flask with 10mL of Dulbecco's Modified Eagle's Medium supplemented with 10% (v/v) fetal bovine serum. After the adherent cells reached 70% confluency, the growth medium was removed and adherent cells scraped and collected. The cells were washed with phosphate buffered saline, centrifuged (500g for 10 min) and then the supernatant was removed, for a total of three washes. The cells were then lysed and protein precipitated using a chloroform:methanol:water extraction (10:10:3, by volume). The precipitated protein was removed, re-solubilized in 4M urea and protein yield quantified with the Bradford protein assay[23].

Preparation of dextran ladder

10mg of Dextran T2000 (Amersham Pharmacia Biotech AB, Uppsala, Sweden) was dissolved in 1mL of 1M TFA and held at 80˚C for 30 minutes. After this partial acid hydrolysis[106], the sample was dried by centrifugal evaporation at room temperature. Once dried, the sample was reduced, desalted and carbon cleaned as described by Jensen *et al*[29] with the following scale-up modifications: reduction was performed with 500µL of 2M $NaBH_4$ in 50mM KOH, the sample was desalted with 100µL packed volume of cation exchange resin and enriched with 100µL packed volume of graphitised carbon. Liquid volumes for desalting and enrichment were increased four-fold from 50µL to 200µ and OMIX C18-100µL pipette tips (Agilent, Santa Clara, CA) were used to contain the increased solid phase resin volumes. The final amount of dextran prepared following graphitised carbon enrichment was quantified using the phenol sulphuric acid method[107].

Glycan release

N- and *O*-glycans were released from 30µg of the protein samples as described by Jensen *et al*[24]. Briefly, the protein samples were spotted on polyvinylidene difluoride membranes (Millipore, Sydney, Australia) and stained with Direct Blue (Sigma-Aldrich, Sydney, Australia). The membrane spots were excised and washed in separate wells in a 96-well plate (Corning Incorporated, NY). *N*-glycans were released using 1U PNGase-F (Promega, Sydney, Australia) at 37°C, overnight. Following *N*-glycan removal, 500mM $NaBH_4$ in 50mM KOH solution was added to immobilised protein spots for 16 h to release reduced *O*-linked glycans by reductive β-elimination.

Released *N*-glycans were reduced and both *N*-glycans and *O*-glycans were desalted and enriched offline using AG 50W-X8 (Bio-Rad, Sydney, Australia) strong cation exchange followed by PGC solid phase extraction micro-columns (Grace, Columbia, MD, USA). Prior to injection, 26ng of the dextran ladder was added to each glycan mixture.

Glycan data acquisition

PGC-LC-ESI-MS/MS experiments were performed on an UltiMate3000 high performance liquid chromatography (HPLC) system (Dionex, Sunnyvale, CA, USA) interfaced with a LTQ Velos Pro ion trap (Thermo Scientific, San Jose, CA, USA). Separations were performed using a PGC LC column (3µm, 100mm x 0.18mm, Hypercarb, Thermo Scientific) maintained at 50°C, for reduced peak width and improve spray stability. Mobile phases used for separation were composed of 10 mM ammonium bicarbonate aqueous solution (solvent A) and 10mM ammonium bicarbonate aqueous solution with 70% (v/v) acetonitrile (solvent B) with a flow rate of 4µl/min. Two gradients were used for normalisation experiments with the first as follows: 0-4.9 min, 1% B; 4.9-50 min, linear increase up to 45.2% B; 50-55 min, held at 99% B; 55-60 min, equilibrated at 1% B for 5 min before next injection - giving a total LC run time of 60 min. The second gradient, simulating reduced glycan retention, was the the same except 0-4.9 min was at 8% B. The gradient used for building the GU library was as follows: 0-4.9 min, 1% B; 5 min, 7.8% B; 5-73 min, 64% B; 73-78 min, 99% B, 78-83 min, equilibrated at 1% B for 5 min before next injection - giving a total LC run time of 83 min.

MS parameters were used as previously optimized and are described by Ashwood *et al*[22] to obtain ten data points per chromatographic glycan peak and MS2 spectra capable of discriminating glycan isomers.

Glycan data analysis

Software-assisted manual glycan identification and annotation of spectra was performed as previously described[22] using Xcalibur (v3.0), RawMeat (v2.1), GlycoMod[74] and GlycoWorkBench (v2.1)[87]. *O*-glycan structural assignment was specifically performed using UniCarb-DB[15] searching of the human mucin dataset created by Jin *et al*[108]. Apex retention time values used were as identified by Skyline (v4.1.1.18151)[109]. Peak area calculation and normalisation to the dextran ladder (also known as global standards) was also performed by Skyline[110][109][109][109][108][108][107][107][106][105][105][105]. Equation determination and relative abundance normalisation was performed with Microsoft Excel (v2010). The glycan library with GU values represented as explicit retention times was built in Excel and a corresponding Skyline library made in “small molecule” mode as previously described by Ashwood *et al*[22].

Spectral library

Raw data files were centroided using msconvert and the SSL file was subsequently populated with appropriate fields from the analysis of raw data files using msaccess with all software obtained from the Proteowizard package[111] (v3.0.10730). Chemical formula and inchikeys were determined using ChemDraw. Unique glycan structural identifiers were assigned to unambiguous glycan structures using GlyTouCan[112] (last accessed: 2018-07-14) and the spectral library constructed using Bibliospec[113]. MS2 spectral annotation was performed using GlycoWorkBench[87] (https://code.google.com/archive/p/glycoworkbench/downloads), manually inspected for verification, and added to the BLIB file using SQLite v3.22 (https://www.sqlite.org/download.html).

Results and discussion

Dextran ladder separated by PGC chromatography

After analysing partially hydrolysed dextran by PGC-LC-ESI-MS/MS, a chemically reduced dextran ladder, composed of subunit lengths varing between 3-13 GU, was detected with $(Glc)_6$ the most abundant subunit (Figure 1A/1C). Dextran ladder subunits less than 3 GU were not detected and the loss of these may be attributed to the graphitised carbon cleanup stage which has been reported to not retain monosaccharides and disaccharides[29]. As dextran ladder subunits greater than 13 GU were not detected, we hypothesise that an acetonitrile composition greater than 50% (v/v) would be required for their elution from the graphitised carbon clean-up. As the PGC column is a stationary phase capable of strongly binding polar analytes[114], we enriched the dextran ladder using an offline carbon cleanup step to limit the ladder range, preventing a loss of column retention strength due to longer dextran subunits remaining on the column with the gradient used.

The elution pattern of the ladder was best described with a logarithmic equation (Figure 1B, Table S-1, $R^2 > 0.99$). Our prepared sample of hydrolysed detran did not change in its measured retention time values with PGC chromatography for 10 technical replicate injections with a median CV of 0.5% for retention time and no observed loss of retention over the 10 runs (Figure S-1). As an assessment of the reproducibility of our PGC-LC-ESI-MS/MS platform for MS1 extracted chromatogram (EIC) based quantitation, we also observed a median CV of 8% for peak area (Figure S-1) based on monoisotopic extracted ion chromatogram (EIC) peak areas.

While the logarithmic elution pattern of a dextran ladder separated on PGC has not previously been reported, 2AB labelled dextran ladder has previously been characterised to elute in a logarithmic pattern using HILIC amide-based columns[95]. On the other hand, a linear elution pattern was observed with permethylated dextran separated on reversed-phase C18 with a linear elution gradient[115]. In addition to research characterising the dextran elution pattern, a labelled dextran ladder separated by HILIC has previously been used to normalise retention time variation[95].

Normalising variation in retention time and MS quantitation

Retention time variation was normalised to an index represented by GU values by converting retention time values using run-specific logarthmic equations based on the dextran ladder. To test the normalisation effectiveness, a batch of 16 samples containing a mixture of paucimannose and high mannose glycans was separated. Then, an artificial retention time shift was introduced by running the last 8 samples with a higher starting organic solvent composition. As shown in Figure 2A, this resulted in a significant disruption in glycan retention times across the entire batch. Instead of using retention time, describing the glycan elution value with GU, obtained through logarithmic transformation of the observed retention times, significantly minimised variation (Figure 2B).

As there are a variety of quantitative methods for glycomics, alternative peak normalisation methods are used depending on the research hypothesis. Despite this, all quantitative glycomics experiments encounter technical variation as a major obstacle for reliable quantitation. An equimolar amount of dextran ladder was added to the previously defined released glycan mixture, as used for the retention time normalisation experiment. Peak area variation of technical replicate samples was reduced by quantifying glycans based on their EIC peak area as a ratio to the internal standards corresponding to the dextran ladder internal standard (Glc4-8)[109]. These subunits of the dextran ladder were chosen for normalisation as they were the most abundant subunits detected and the use

of multiple standards across the chromatographic run has been demonstrated to be a more robust than normalisation to a single standard[116].

As shown in Figure 2C, the variability between glycan peak areas across technical replicates was measured to assess post-acquisition normalisation methods. Without normalisation, the median co-efficient of variation (CV) in glycan peak areas was 17%. Relative abundance, a normalisation method which quantifies individual glycan structures based on their EIC peak area compared to the total EIC peak area for all detected glycans, is typically used in LC-MS-based glycomics[29], and more than halved this variation (6% median CV). Quantifying glycans based on their peak areas ratios relative to the dextran ladder standards was performed in parallel and found to reduce variation to the same extent (6% median CV).

Relative abundance quantitation is frequently performed when analysing quantitative differences in glycan populations as it does not require internal standards. This form of quantitation also provides normalisation of technical variation due to each glycan being a percentage of the total glycan abundance in a single analytical run[44]. However one inherent limitation in this approach is the inability to detect overall perturbations in protein glycosylation which change the total amount of glycosylation such as inhibition of the *N*-glycosylation pathway[117]. With the approach described here, normalisation to internal standards added in equal amount to all samples, these pertubations can be monitored[118].

We have normalised both glycan peak areas and retention time using Skyline, an open-source software package for LC-MS experiments[109]. The use of software instead of manual analysis is desireable for automated glycan profiling as it results in improved reproducibility and higher throughput[119]. While Skyline is currently limited to retention time normalisation using linear equations, normalisation of glycan retention time was still suitable with five point calibration (Figure S-2, R^2 =0.998) however assignment of glycan GU values was performed manually using seven point logarithmic calibration for improved accuracy (R^2 >0.999).

Normalisation of glycan retention time to GU values allows the observed GU value of a glycan to be a widely applicable parameter in structure assignment. Following characterisation of the glycan structures using orthogonal MS/MS annotation, a GU library with assigned glycan structures was constructed (File S-2). An approach such as this, primarily using glycosidase treatment, has been successfully applied for HILIC-based labelled-glycan separation[95] but does not yet exist for PGC-based and/or unlabelled glycan separation.

Development of a glycan GU library for PGC-based chromatography

Thirty micrograms of purified glycoproteins, including lactoferrin, fetuin and ribonuclease B, carrying the main classes of *N*-glycans (high mannose, hybrid, paucimannose, complex) and *O*-glycans (*O*-GalNAc multiple core-types and *O*-Mannose) had their glycans released, combined with the purified dextran ladder and analysed with PGC-LC-ESI-MS/MS. As a result, 189 *N*-glycan and 100 *O*-glycan unique structures were assigned GU values (Table 1 and File S-2). The range of our prepared dextran ladder was suitable for assignment of GU values for *N*-glycans from one of the smallest and most hydrophilic glycans (M3B, 4 GU) to core-fucosylated tetra-antennary glycans (A4G(4)4F1S4, 13 GU) having GU values within the range of our ladder. Several *O*-glycan structures were outside this range with 5 structures assigned a GU value less than 3 and 4 structures with assigned GU values greater than 13 (Table S-2). With the logarithmic elution behaviour of the dextran ladder, extrapolated GU values for glycans outside 3-13 GU can serve as a suitable strategy until further improvements allows detection of an increased dextran ladder length on PGC-LC-ESI-MS/MS.

To ensure no duplication in the library, glycans with precursor values within *m/z* ±0.5 and retention values within ±0.2 GU between glycan library entries were excluded. More stringent filtering could be performed with longer gradients which would give more accurate retention times. In addition to glycan structures (represented by GlyTouCan accession values[112]) assigned to GU value, other important metadata such as observed and theoretical *m/z* values and MIRAGE compliant parameters are included in the supplementary files, allowing this library to be used with other LC-MS-based glycan analysis platforms[16].

To evaluate the effectiveness of our library for analysing biologically relevant glycan mixtures, P*N*Gase-F released *N*-glycans from 30 µg of a protein extract from U87MG cell line were mixed with the dextran ladder and analysed by PGC-LC-ESI-MS/MS. Filtering for unique structures, using derived GU values and precursor masses, 71 glycan structures were matched to the standards in the library, utilising 1/3rd of our mammalian GU library (Figure 3A and B).

The data were analysed manually in parallel and 46 *N*-glycan structures not included in the initial GU library were characterised and identified in the cell lysate. These *N*-glycan structures were predominantly complex tri-antennary and tetra-antennary *N*-glycans which did not exist in our original library due to different sialic acid linkages and/or the presence of the LacDiNAc epitope which were evidently not detected from our standard glycoproteins (Figure 3C). These new structures were subsequently added to the library to completely cover the *N*-glycans detected.

While GU values obtained using PGC chromatography appear to be suitable for discriminating between glycan isomers (e.g. glycan structures differing in sialic acid linkage), the additional use of diagnostic ions can further increase confidence in glycan structure assignment[22]. This workflow provides the first PGC-based GU library for automated glycan characterisation by LC-MS and as a result of conversion of retention time to GU values, the library is applicable for PGC-based LC separations irrespective of system and column condition.

Understanding PGC-based glycan separation

The assignment of GU values to a diverse range of glycan structures separated on PGC affords an opportunity to further understand the mechanism by which these glycan structures are separated. The derived GU values for *N*-glycans with and without substructural features such as core-fucose and bisecting GlcNAc were compared and strong correlative trends were observed (Figure 4, R^2 >0.9).

For *N*-glycans, the presence of bisecting GlcNAc resulted in reduced retention with the magnitude depending on the retention characteristics of the non-bisected precursor glycan structure. For one of the earlier eluting glycan structures (M3, 5.5 GU) the bisected version elutes 2.1 GU earlier (M3B, 3.4 GU) whereas for a later eluting structure (F(6)A1, 6.9 GU) the bisected version elutes 1.2 GU earlier (F(6)A1B, 5.7 GU). An inverse relationship in GU shifts was observed for the addition of core-fucose which resulted in an increase in retention, the magnitude of which, depended again on the precursor non-fucoslyated glycan structure. For one of the earlier eluting glycan structures (A1B, 3.6 GU) the core-fucosylated form elutes 1.6 GU later (F(6)A1B, 5.2 GU) whereas for a later eluting structure (M3, 6.3 GU) the core-fucosylated version elutes 2.2 GU later (M3F, 8.5 GU).

The desire to describe the mechanism by which PGC separates glycan structures has been a goal for several groups in the past[41, 98, 100]. Our GU library is consistent with these PGC elution rules and provides additional detail by providing predictive GU models, enabled as a result of our conversion of retention time to a system-independent GU values. As this library becomes further populated with a

wider range of glycan structures, the net retention effect of other sub-structural elements, such as sialic acid linkages, could be modelled in the same way as demonstrated here.

Development of a dextran ladder spectral library

To extend the applicability of the dextran ladder towards resolving mixtures containing isomers of the hydrolysed dextran, we developed a spectral library of the dextran ladder composed of 3-13 units. Our dextran ladder mixed with *O*-mannose glycans from cellobiohydrolase I, a fungal protein secreted by *T. reesei*, was chosen as a challenging sample to test our spectral library approach due to the glycans being isomeric with the dextran ladder[23]. As shown in Figure 5A, three polysaccharides were detected with masses consistent with a composition of $(Hex)_4$ (*m/z* 667.23 ±0.35) and, as multiple MS2 scans were collected for each isomer, extracted ion chromatograms for each theoretically expected product ion were contructed (Figure 5B).

Following analysis of only the dextran ladder, a spectral library was populated with observed and annotated MS2 spectra with one example shown in Figure 5C. The use of the spectral matching feature in Skyline identified the peak at 19.9 min to be the most probable match out of the three glycan peaks. The strong match between the peak at 19.9 minutes (dotp 0.92) and the spectral library (Figure 5C) is due to a single prominent peak at *m/z* 383.1 representing a cross-ring fragment of a glucose tetramer which is specific for dextran rather than the isomeric *O*-mannosylated glycans[120]. The same correct peak picking was achieved for all other isomers ($(Hex)_3$, $(Hex)_4$, $(Hex)_5$ and $(Hex)_6$) due to the characteristic cross-ring fragmentation pattern of the glucose-based dextran (Figure S-3).

The assigned dotP score is as defined by Toprak *et al*[121], using a normalised spectral constrast angle comparison of library and unknown spectra which has previously been used for targeted peptide spectral matching. The normalised spectral contrast angle comparison was found to sensitively quantify the similarity of spectra and, although we were analysing glycans instead of peptides, serves as a useful foundation for future studies in spectral matching methods. Our development of an annotated spectral library, combined with the unique fragmentation pattern of dextran, allows automated assignment of the dextran ladder subunits to discrete retention times without requiring manual peak selection.

Data availability and supporting information

In order to adhere to the MIRAGE guidelines[40], complete LC-MS method parameters and data links between raw file names, corresponding Skyline assays and glycan GU libraries listed in file S-1 as well as all supplementary figures and tables. The glycan GU libraries are available in file S-2. The raw data files were uploaded to Chorus (Project 1399). Skyline assays, glycan GU libraries and the annotated dextran ladder spectral library can be found on Panorama (panoramaweb.org/project/__r4143/begin.view?). Glycan MS2 spectra annotated by GlycoWorkBench are found in file S-3.

Conclusions

These results demonstrate the first example of system-independent assignment of glycan structure based only on precursor mass and GU value using PGC-based LC-MS, with orthogonal confirmation by MS2. A glycan library was constructed and our described method was applied to *N*-glycans released from a U87MG cell line. Despite the identification of peaks with unassigned glycan structures, parallel software-assisted manual analysis allowed extension of the library for complete

N-glycan assignment of those detected from the cell line. Arising from current technical limitations regarding exact glycan structure determination, some assigned glycan structures are partially ambiguous but continuing efforts in glycan characterisation by alternative MS/MS modes[122] and ion-mobility/spectroscopy[123] could overcome this limitation for complete glycan structural characterisation.

Although this library is by no means comprehensive for all glycan structures and therefore applicable towards all sample types, it serves as a foundation for further extension into a comprehensive glycan library.

References

(1) Rudd, P. M.; Rudan, I.; Wright, A. F. High-Throughput Glycome Analysis Is Set to Join High-Throughput Genomics. J. Proteome Res. 2009, 8 (3), 1105.

(2) Everest-Dass, A. V; Moh, E. S. X.; Ashwood, C.; Shathili, A. M. M.; Packer, N. H. Human Disease Glycomics: Technology Advances Enabling Protein Glycosylation Analysis – Part 1. Expert Rev. Proteomics 2017, 0 (0), 14789450.2018.1421946.

(3) Royle, L.; Campbell, M. P.; Radcliffe, C. M.; White, D. M.; Harvey, D. J.; Abrahams, J. L.; Kim, Y. G.; Henry, G. W.; Shadick, N. A.; Weinblatt, M. E.; et al. HPLC-Based Analysis of Serum N-Glycans on a 96-Well Plate Platform with Dedicated Database Software. Anal. Biochem. 2008, 376 (1), 1–12.

(4) Jensen, P. H.; Karlsson, N. G.; Kolarich, D.; Packer, N. H. Structural Analysis of N- and O-Glycans Released from Glycoproteins. Nat. Protoc. 2012, 7, 1299–1310.

(5) Melmer, M.; Stangler, T.; Premstaller, A.; Lindner, W. Comparison of Hydrophilic-Interaction, Reversed-Phase and Porous Graphitic Carbon Chromatography for Glycan Analysis. J. Chromatogr. A 2011, 1218 (1), 118–123.

(6) Pabst, M.; Bondili, J. S.; Stadlmann, J.; Mach, L.; Altmann, F. Mass + Retention Time = Structure: A Strategy for the Analysis of N-Glycans by Carbon LC-ESI-MS and Its Application to Fibrin N-Glycans. Anal. Chem. 2007, 79, 5051–5057.

(7) Pabst, M.; Altmann, F. Influence of Electrosorption, Solvent, Temperature, and Ion Polarity on the Performance of LC-ESI-MS Using Graphitic Carbon for Acidic Oligosaccharides. Anal. Chem. 2008, 80, 7534–7542.

(8) Aldredge, D.; An, H. J.; Tang, N.; Waddell, K.; Lebrilla, C. B. Annotation of a Serum N-Glycan Library for Rapid Identification of Structures. J. Proteome Res. 2012, 11 (3), 1958–1968.

(9) Abrahams, J. L.; Campbell, M. P.; Packer, N. H. Building a PGC-LC-MS N-Glycan Retention Library and Elution Mapping Resource. Glycoconj. J. 2018, 35 (1), 15–29.

(10) Guile, G. R.; Rudd, P. M.; Wing, D. R.; Prime, S. B.; Dwek, R. A. A Rapid High-Resolution High-Performance Liquid Chromatographic Method for Separating Glycan Mixtures and Analyzing Oligosaccharide Profiles. Anal. Biochem. 1996, 240 (2), 210–226.

(11) Luo, Q.; Rejtar, T.; Wu, S.-L.; Karger, B. L. Hydrophilic Interaction 10 Mm I.D. Porous Layer Open Tubular Columns for Ultratrace Glycan Analysis by Liquid Chromatography-Mass Spectrometry. J Chromatogr. A 2009, 1216 (8), 1223–1231.

(12) Ahn, J.; Bones, J.; Yu, Y. Q.; Rudd, P. M.; Gilar, M. Separation of 2-Aminobenzamide Labeled Glycans Using Hydrophilic Interaction Chromatography Columns Packed with 1.7 Mm Sorbent. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 2010, 878 (3–4), 403–408.

(13) Stockmann, H.; Adamczyk, B.; Hayes, J.; Rudd, P. M. Automated, High-Throughput IgG-Antibody Glycoprofiling Platform. Anal. Chem. 2013, 85 (18), 8841–8849.

(14) Bundle, D. R.; Jennings, H. J.; Kenny, C. P. Studies on the Group-Specific Polysaccharide of Neisseria Meningitidis Serogroup X and an Improved Procedure for Its Isolation. J. Biol. Chem. 1974, 249 (15), 4797–4801.

(15) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. GlycoBase and AutoGU: Tools for HPLC-Based Glycan Analysis. Bioinformatics 2008, 24 (9), 1214–1216.

(16) Ashwood, C.; Abrahams, J. L.; Nevalainen, H.; Packer, N. H. Enhancing Structural Characterisation of Glucuronidated O -Linked Glycans Using Negative Mode Ion Trap-HCD Mass Spectrometry. Rapid Commun. Mass Spectrom. 2017, 31, 851–858.

(17) Schneider, P.; Ferguson, M. A. J. Microscale Analysis of Glycosylphosphatidylinositol Structures. Methods Enzymol. 1995, 250 (C), 614–630.

(18) Masuko, T.; Minami, A.; Iwasaki, N.; Majima, T.; Nishimura, S. I.; Lee, Y. C. Carbohydrate Analysis by a Phenol-Sulfuric Acid Method in Microplate Format. Anal. Biochem. 2005, 339 (1), 69–72.

(19) Ashwood, C.; Lin, C.-H.; Thaysen-Andersen, M.; Packer, N. H. Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. J. Am. Soc. Mass Spectrom. 2018.

(20) Cooper, C. A.; Gasteiger, E.; Packer, N. H. GlycoMod--a Software Tool for Determining Glycosylation Compositions from Mass Spectrometric Data. Proteomics 2001, 1, 340–349.

(21) Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M. GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. J. Proteome Res. 2008, 7, 1650–1659.

(22) Hayes, C. A.; Karlsson, N. G.; Struwe, W. B.; Lisacek, F.; Rudd, P. M.; Packer, N. H.; Campbell, M. P. UniCarb-DB: A Database Resource for Glycomic Discovery. Bioinformatics 2011, 27 (9), 1343–1344.

(23) Jin, C.; Kenny, D. T.; Skoog, E. C.; Padra, M.; Adamczyk, B.; Vitizeva, V.; Thorell, A.; Venkatakrishnan, V.; Lindén, S. K.; Karlsson, N. G. Structural Diversity of Human Gastric Mucin Glycans. Mol. Cell. Proteomics 2017, mcp.M117.067983.

(24) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. Bioinformatics 2010, 26 (7), 966–968.

(25) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. Bioinformatics 2008, 24 (21), 2534–2536.

(26) Tiemeyer, M.; Aoki, K.; Paulson, J.; Cummings, R. D.; York, W. S.; Karlsson, N. G.; Lisacek, F.; Packer, N. H.; Campbell, M. P.; Aoki, N. P.; et al. GlyTouCan: An Accessible Glycan Structure Repository. Glycobiology 2017, 27 (10), 915–919.

(27) Frewen, B.; MacCoss, M. J. Using BiblioSpec for Creating and Searching Tandem MS Peptide Libraries. Curr. Protoc. Bioinforma. 2007, 1–12.

(28) Forgács, E.; Cserháti, T. Relationship between Retention Characteristics and Physicochemical Parameters of Solutes on Porous Graphitized Carbon Column. J. Pharm. Biomed. Anal. 1998, 18, 505–510.

(29) Zhou, S.; Hu, Y.; Mechref, Y. High-Temperature LC-MS/MS of Permethylated Glycans Derived from Glycoproteins. Electrophoresis 2016, 37 (11), 1506–1513.

(30) Mizuno, H.; Ueda, K.; Kobayashi, Y.; Tsuyama, N.; Todoroki, K.; Min, J. Z.; Toyo'oka, T. The Great Importance of Normalization of LC–MS Data for Highly-Accurate Non-Targeted Metabolomics. Biomed. Chromatogr. 2017, 31 (1), 1–7.

(31) Moh, E. S. X.; Thaysen-Andersen, M.; Packer, N. H. Relative versus Absolute Quantitation in Disease Glycomics. Proteomics - Clin. Appl. 2015, 9 (3–4), 368–382.

(32) Elbein, A. Glycosylation Inhibitors for N-Linked Glycoproteins. In Methods in Enzymology; 1987; Vol. 138, pp 661–709.

(33) Zaia, J. Mass Spectrometry and the Emerging Field of Glycomics. Chemistry and Biology. 2008, pp 881–892.

(34) Shubhakar, A.; Reiding, K. R.; Gardner, R. A.; Spencer, D. I. R.; Fernandes, D. L.; Wuhrer, M. High-Throughput Analysis and Automation for Glycomics Studies. Chromatographia 2014, 78 (5–6), 321–333.

(35) Struwe, W. B.; Agravat, S.; Aoki-Kinoshita, K. F.; Campbell, M. P.; Costello, C. E.; Dell, A.; Feizi, T.; Haslam, S. M.; Karlsson, N. G.; Khoo, K. H.; et al. The Minimum Information Required for a Glycomics Experiment (MIRAGE) Project: Sample Preparation Guidelines for Reliable Reporting of Glycomics Datasets. Glycobiology 2016, 26 (9), 907–910.

(36) Čmelík, R.; Štikarovská, M.; Chmelík, J. Different Behavior of Dextrans in Positive-Ion and Negative-Ion Mass Spectrometry. J. Mass Spectrom. 2004, 39 (12), 1467–1473.

(37) Toprak, U. H.; Gillet, L. C.; Maiolica, A.; Navarro, P.; Leitner, A.; Aebersold, R. Conserved Peptide Fragmentation as a Benchmarking Tool for Mass Spectrometers and a Discriminating Feature for Targeted Proteomics. Mol. Cell. Proteomics 2014, 13 (8), 2056–2071.

(38) Tang, Y.; Wei, J.; Costello, C. E.; Lin, C. Characterization of Isomeric Glycans by Reversed Phase Liquid Chromatography-Electronic Excitation Dissociation Tandem Mass Spectrometry. J. Am. Soc. Mass Spectrom. 2018, 1–13.

(39) Manz, C.; Pagel, K. Glycan Analysis by Ion Mobility-Mass Spectrometry and Gas-Phase Spectroscopy. Curr. Opin. Chem. Biol. 2018, 42, 16–24.

(40) Hayase, T.; Lee, Y. C.; Rice, K. G.; Dziegielewska, K. M.; Mark, K.; Thomas, R. Comparison of N-Glycosides of Fetuins from Different Species and Human A2-HS-Glycoprotein. Biochemistry 1992, 31 (20), 4915–4921.

(41) Spik, G.; Coddeville, B.; Mazurier, J.; Bourne, Y.; Cambillaut, C.; Montreuil, J. Primary and Three-Dimensional Structure of Lactotransferrin (Lactoferrin) Glycans. In Lactoferrin Structure and Function; 1994; pp 21–33.

(42) Loke, I.; Østergaard, O.; Heegaard, N. H. H.; Packer, N. H.; Thaysen-Andersen, M. Paucimannose-Rich N -Glycosylation of Spatiotemporally Regulated Human Neutrophil Elastase Modulates Its Immune Functions. Mol. Cell. Proteomics 2017, 16 (8), 1507–1527.

(43) Pučić, M.; Knežević, A.; Vidič, J.; Adamczyk, B.; Novokmet, M.; Polašek, O.; Gornik, O.; Šupraha-Goreta, S.; Wormald, M. R.; Redžić, I.; et al. High Throughput Isolation and Glycosylation Analysis of IgG–Variability and Heritability of the IgG Glycome in Three Isolated Human Populations. Mol. Cell. Proteomics 2011, 10 (10), M111.010090.

(44) Mattu, T. S.; Pleass, R. J.; Willis, A. C.; Kilian, M.; Wormald, M. R.; Lellouch, A. C.; Rudd, P. M.; Woof, J. M.; Dwek, R. A. The Glycosylation and Structure of Human Serum IgA1, Fab, and Fc Regions and the Role of N-Glycosylation on Fcα Receptor Interactions. J. Biol. Chem. 1998, 273 (4), 2260–2272.

(45) Hui, J. P. M.; Lanthier, P.; White, T. C.; McHugh, S. G.; Yaguchi, M.; Roy, R.; Thibault, P. Characterization of Cellobiohydrolase I (Cel7A) Glycoforms from Extracts of Trichoderma Reesei Using Capillary Isoelectric Focusing and Electrospray Mass Spectrometry. J. Chromatogr. B Biomed. Sci. Appl. 2001, 752, 349–368.

(46) Thomsson, K. A.; Karlsson, H.; Hansson, G. C. Sequencing of Sulfated Oligosaccharides from Mucins by Liquid Chromatography and Electrospray Ionization Tandem Mass Spectrometry. Anal. Chem. 2000, 72 (19), 4543–4549.

Acknowledgements

This work was supported by the Australian Research Council (CE140100003).

Figures and Tables

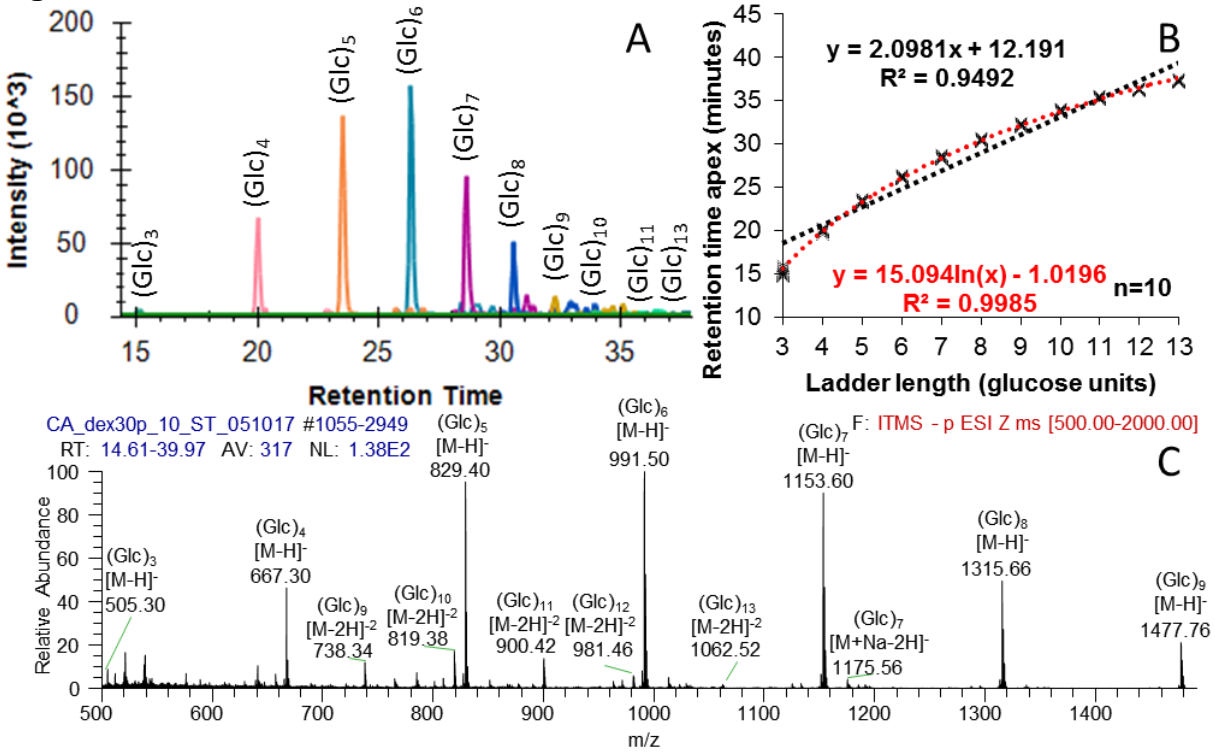


Figure 1 Characterisation of the dextran ladder using PGC-LC-ESI-MS/MS. **Panel A** Combined extracted ion chromatograms of each observed subunit of the dextran ladder. **Panel B** Best equation fitting of retention time for GU assignment (110 data points) with linear equation in black and logarithmic equation in red. **Panel C** Summed MS spectra across the dextran ladder retention time range

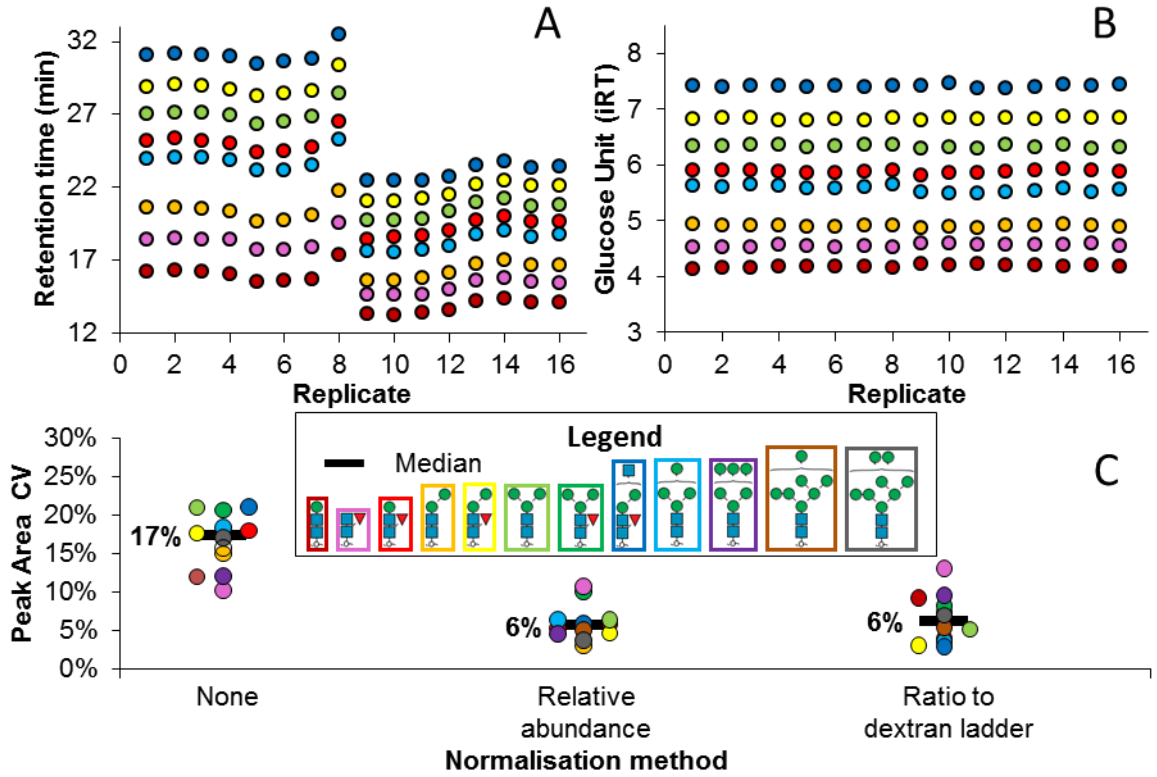


Figure 2 Normalising two aspects of LC-MS variation using a dextran ladder. Glycan structure colour legend corresponds to all panels. **Panel A** Observed retention times of a glycan mixture over 16 technical replicates with a simulated loss of retention for replicates 9-16. **Panel B** Normalised retention times of the 16 technical replicates shown in panel A. **Panel C** Peak area variation comparison of post-acquisition normalisation techniques for glycan quantitation.

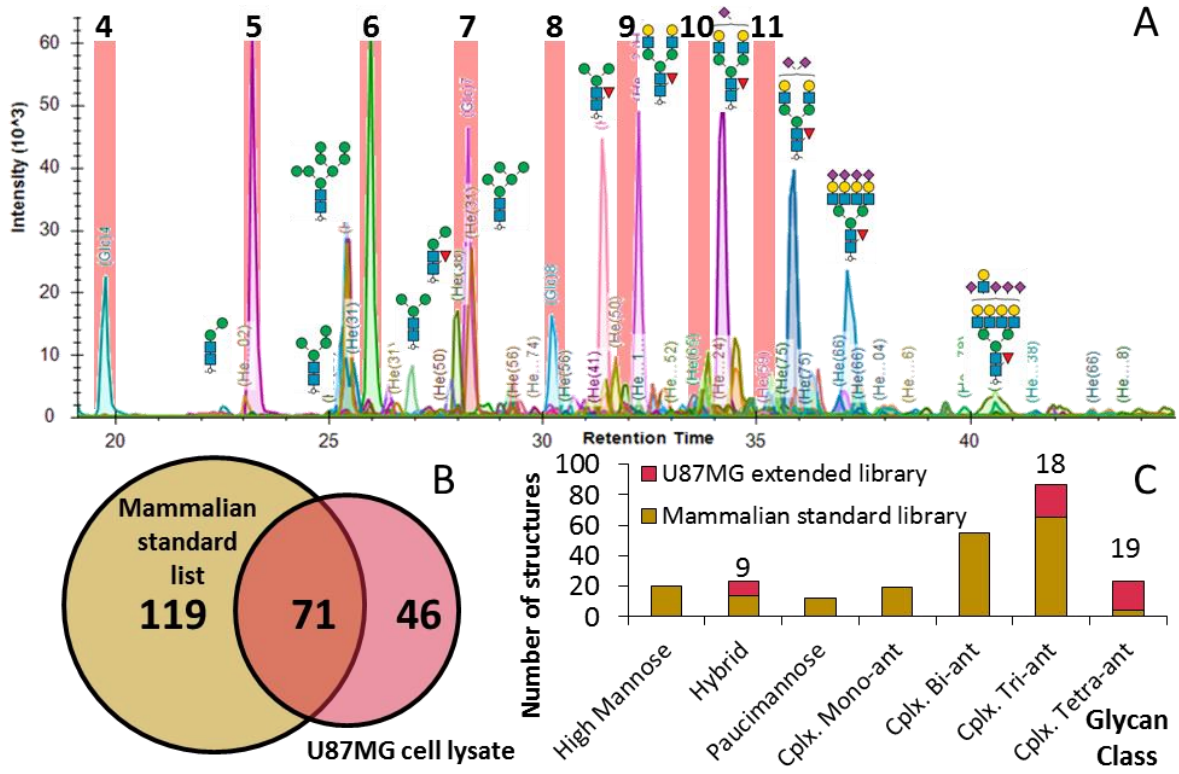


Figure 3 Application of a PGC-based GU library for identification of *N*-glycans released from a U87MG cell lysate. **Panel A** Extracted ion chromatograms of U87MG glycans mixed with the dextran ladder, only dextran subunits 4-11 highlighted. **Panel B** Coverage of the GU library comprising glycans from

purified glycoprotein standards. **Panel C** Comparison of glycan class coverage following extension of GU library with cell lysate *N*-glycans

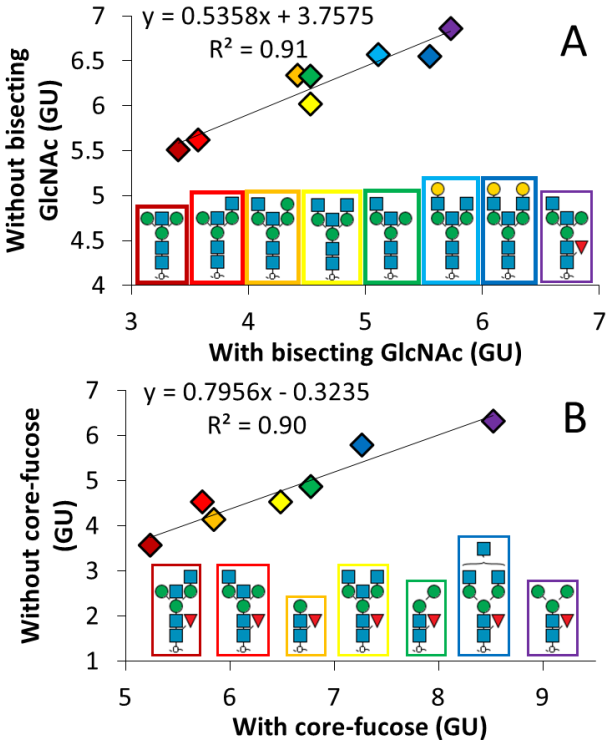


Figure 4 Identification of equations to predict glycan GU values by correlation of base structure GU value with **A)** bisected and **B)** core-fucosylated glycan structures

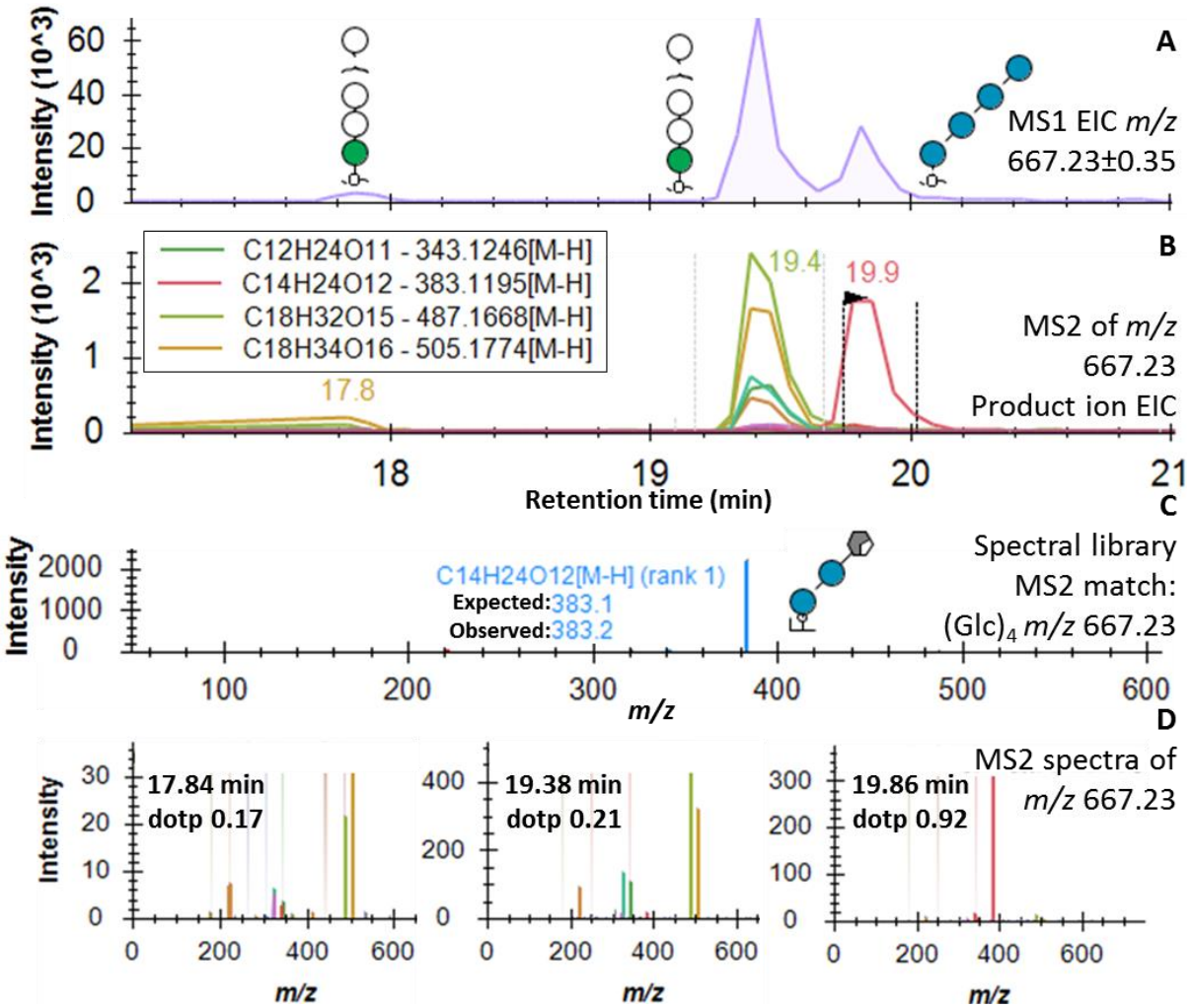


Figure 5 Automated glucose-based dextran ladder peak assignment in a glycan sample containing isomeric mannose-based glycans using a spectral library approach. **Panel A** MS1 extracted ion chromatograms for the glycan composition $(Hex)_4$ *m/z* 667.2302 $[M-H]^-$, structures identified. **Panel B** MS2 extracted ion chromatograms of abundant product ions for the glycan composition defined in panel A. **Panel C** Annotated MS2 spectra of dextran standard, $(Glc)_4$ *m/z* 667.2302 $[M-H]^-$ for use in spectral matching. **Panel D** Isomeric $(Hex)_4$ MS2 spectra with spectral similarity score shown as dotp value.

Standard (Protein or glycolipid)	Number of glycan structures	Main type of glycan
Bovine fetuin (*N*-glycans)	17	Complex Sialylated[124]
Bovine lactoferrin (*N*-glycans)	72	Complex Fucosylated[125]
Human neutrophil elastase (*N*-glycans)	16	Paucimannose[126]
Human IgG (*N*-glycans)	6	Complex Non-sialylated[127]
Human IgA (*N*-glycans)	59	Complex Sialylated, Hybrid and Bisected[128]
Human lactoferrin (*N*-glycans)	17	Complex Fucosylated[125]
Fungal cellobiohydrolase I (*N*-glycans)	2	High mannose[129]
Total *N*-glycan unique mammalian structures	189	All types
Porcine gastric mucin (*O*-glycans)	92	*O*-GalNAc, multiple core types[130]
Fungal Cellobiohydrolase I (*O*-glycans)	8	*O*-mannosylation[23]

Table 1 Purified glycoproteins used for GU library construction of released glycans

SECTION 2 – INVESTIGATING THE TLR4 COMPLEX AND EFFECT OF ITS ACTIVATION ON PROTEIN N-GLYCOSYLATION

Preface

While section 1 of this thesis has focused on optimisation of data acquisition and data analysis to improve the quality of data collected to study protein glycosylation, this thesis section is the application of these techniques, and other approaches, to investigate a challenging area of immunological research: chronic pain.

Introduction

The TLR4 signalling complex as a mediator of disease states

Pattern recognition receptors, including toll-like receptors, are responsible for the immediate host defence system response to recognise specific foreign and host biomolecules. These foreign and host biomolecules are described as damage-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs), respectively[131]. Through TLR4 binding to a molecular pattern such as lipopolysaccharide (LPS), a classic PAMP, a complex signalling process is initiated which results in induction of cytokines and other inflammatory mediators[132, 133]. Despite this being a convoluted signalling process, activation of the TLR4 signalling pathway is considered immediate (reported cell response within 2 hours[134]) which contrasts with other immune responses such as antibody and T-cell activation.

As the TLR4 signalling pathway is a contributing cause for a wide range of human diseases with complex mechanisms including diabetes[135, 136], alcoholism [137, 138] and myocardial ischemia[139], expression of the TLR4 complex has been mapped to specific-tissue regions and cell-types. While TLR4 mRNA has been detected in almost all tissues, antibody detection of TLR4 has revealed its expression in the brain, endocrine tissues, bone marrow and immune system, and the gastrointestinal tract[140]. The brain has been one area of focus in particular for understanding the role of TLR4 in recognising DAMPs and PAMPs involved in chronic pain[137, 141–143]. Further complicating research into the TLR4 pathway of brain, there are various cell types present, and each cell type can play unique roles in the overall immune response to neuropathic pain[144].

The TLR4 signalling pathway

Toll-like receptors are trans-membrane receptors that, upon binding to ligands, initiate intracellular signalling pathways. These toll-like receptors which respond to gram-positive bacterial or fungal

infection were first discovered in *Drosophila melanogaster* [145]**.** Similarly to humans, activation of toll-like receptors in *Drosophila* results in the activation of nuclear factor kappa B (NF-KB) [146].

In humans, 13 toll-like receptors have currently been identified but their significance was not understood until mutations of a gene with unknown protein encoding in mice were performed, identifying defective immune responses to LPS[147]. Through subsequent genetic and physical mapping of this gene region, Toll-like receptor 4 (TLR4) was identified as the candidate gene responsible for this defective response to LPS[148].

For toll-like receptors generally, ligand binding induces receptor dimerization and additional conformational changes, attracting adaptor molecules. The adaptor molecules, including Myeloid Differentiation factor 88 (MyD88), toll-interleukin 1 receptor (TIR) domain-containing adaptor protein (TIRAP) and TIR-domain-containing adapter-inducing interferon-β (TRIF), in turn stimulate NF-κB which in turn, activates proinflammatory cytokines. These proinflammatory cytokines, TNFα, IL-1β, IL-6, IL-18, have been understood to influence chronic pain states and other immune responses[149].

For TLR4 specifically, its activation can be mediated by two adaptor molecules, TIRAP and TRAM, which can activate two major pathways: the TIRAP-MYD88 pathway, for early NF-KB activation and inflammatory cytokine production, and the TRIF- toll-like receptor adaptor molecule 1 (TRAM) pathway, which up-regulates genes encoding type I interferons, and activates late phase NF-KB through secreted TNFα[150]. MyD88 is the predominant signalling pathway for the TLR4-mediated LPS response[151].

The initial activation of the TLR4 protein complex is of interest, as it is frequently a target of pharmacological efforts to prevent potentially fatal cytokine storms resulting from repeated TLR4 activation[152]. For LPS stimulation specifically, liposaccharide binding protein (LBP) and CD14 mediate the binding of LPS to MD-2, a secreted protein. This MD-2/LPS complex subsequently acts as a ligand for TLR4, starting the intracellular signalling pathway[153]. To study the proteins part of the TLR4 complex, TLR4, MD-2 and CD14, crystallography has been the predominant technique in the context of their inactivation by small molecule inhibitors.

Models to study TLR4 activation

Toll-like receptors are expressed in a variety of mammalian immune system-related cell types including B cells, natural killer cells and macrophages, as well as non-immune cells such as epithelial and endothelial cells[140]. In the brain, neurons were originally implicated as the sole mediators in pain, however the roles of other cell types such as microglia and astrocytes in pain mediation are steadily being realised[154, 155] as being partly responsible for the conception and modulation of physiological pain states[155, 156].

Looking for avenues to modulate these pain states through pharmacological action, a variety of TLR4 signalling pathway inhibitors have been developed which target several proteins involved in the complex namely, TLR4 and MD-2. These inhibitors include lipid-based molecules (Eritoran[157]) and smaller molecules (Resatorvid[158] and a derivative of 2-acetamidopyranoside[159]) and these antagonists have been demonstrated to prevent[160] and reverse preclinical models of neuropathic pain[161].

The majority of studies investigating TLR4 activation use LPS stimulation with cell lines as a model for the response to PAMPs and DAMPs with a typical focus on CD14, MD2 and TLR4 as a complex[162–166]. To study TLR4 activation in animals, mouse or rats are typically chosen as model organisms. Instead of the LPS administration which is used in the cell line models, TLR4 activation is typically induced through unilateral sciatic nerve chronic constriction injury (CCI) which develops into neuropathic pain[160, 161, 167]. This CCI method has been established to produce disorders of pain sensation similar to those seen in humans, thus serving as a model system to investigate chronic neuropathic pain[168, 169].

The neuropathic pain generated as part of this model is then monitored by behavioural measurements including mechanical withdrawal threshold and thermal withdrawal latency[168]. As pain transmission from the peripheral nervous system to the central nervous system can be modulated, the animal response to pain can be reduced, unmodified or amplified which can be determined using these behavioural measurements[170, 171], with an amplified animal response to pain known clinically as hyperalgesia. In addition to these forms of modulation, allodynia, heightened sensitivity to typically non-painful stimuli, can also occur[172]. The brain is responsible for the processing of these pain states with modulation of these pain signals occurring in the spinal cord dorsal horns[172]. While the role of cells and proteins of the TLR4-complex have been realised for transmission of pain signals from the peripheral nervous system to the brain, the role of protein glycosylation in cell to cell signalling and subsequent signal transduction has not been understood in the neuropathic pain area.

CHAPTER 7 – CHARACTERISING PROTEIN *N*-GLYCOSYLATION CHANGES FOLLOWING TLR4 ACTIVATION

7.1 PREFACE

Following the development of improved methods for analysing protein glycans as described in Section1 of this thesis, we sought to investigate the effect on protein glycosylation produced by TLR4 activation. As identified in Paper 2, the role of protein glycosylation has not been studied in any pain models and is further compounded by the lack of quantitative glycan structural analysis in this area. Quantitative analysis of protein glycosylation could serve as a useful foundation for understanding the role of protein glycosylation in the central nervous system inflammatory response. Furthermore, as protein glycosylation is a mediator of cell signalling and communication[173–175], glycans have been demonstrated to play a role in several facets the human inflammatory response[176–178]. In this chapter, we aimed to provide a first glimpse into the effect TLR4 activation has on cellular and secreted protein *N*-glycosylation.

Overall, we aim to evaluate protein *N*-glycosylation changes in the components of TLR4 pathway activation both *in vitro* and *in vivo*. Such changes would be expected to produce a different cell surface architecture. This overall aim includes four sub aims:

- to determine whether *N*-glycan structures differ on proteins secreted into the cell culture media *in vitro* from cell lines following TLR4 pathway activation by LPS stimulation;
- to investigate whether protein *N*-glycan structure profiles vary in different mouse brain regions, and whether there are glycomic changes in those regions corresponding to the animal behaviour arising from TLR4 pathway activation by neuropathic pain *in vivo*;
- to measure any changes in protein *N*-glycan structure profiles of different mouse spinal cord regions (L3,4,5) responding to TLR4 pathway activation by neuropathic pain *in vivo*;
- to investigate whether unilateral CCI to the mouse *in vivo,* produces spinal cord side-specific *N*-glycan structure differences.

7.2 Results and Discussion

Methods for the results generated in this section can be found described in detail in Chapter 2.

7.2.1 Secreted protein glycosylation in TLR4-activated cell lines

Secreted proteins are a useful starting point for investigating the cellular response to inflammation as these proteins often act as co-factors or adaptor molecules for membrane-based receptors (such as MD-2 and secreted CD14)[179, 180]. To examine the effect of TLR4 activation on secreted protein glycosylation, a human glioblastoma cell line, U87MG, was treated with LPS, a known agonist of the TLR4 pathway. A HEK293 cell line, known to not express protein components of the TLR4 pathway [181], was also treated with LPS to act as a negative control. Before the addition of LPS and commencement of determining the secreted protein glycan changes over time, the cells were conditioned in serum-free media for 24 hours at 70% confluence. Three washes with serum-free media was performed prior to 24-hour conditioning to ensure no carryover of serum glycoproteins and the 24 hour conditioning was used to allow proteins to be secreted from the cell line and therefore obtain a protein N-glycosylation profile at time point 0.

As time-course studies investigating the effect of TLR4 activation on secreted protein glycosylation have not been performed previously, a wide range of time points (0-44 hours) was chosen after LPS administration. The use of serum-free media did not result in greatly reduced cell viability for up to 57 hours after removal of serum (85% average cell viability, Supp Table 7.1) with a 10% decrease in cell viability of U87MG cells at 68 hours following serum-free media replacement. The HEK293 cells had a consistent cell viability of 89% up to and including 68 hours after serum-free media replacement. At each time point up to and including 44 hours after LPS administration, secreted proteins were collected from the serum-free culture media and proteins extracted using acetone precipitation. The glycans attached to these secreted proteins were subsequently released and then analysed with PGC-LC-ESI-MS/MS.

For both human cell lines, fifty-six *N*-glycan structures that were identified and quantified across the time points were released from the secreted proteins (Supp Table 7.2). The structures comprised the main glycan classes (high mannose, hybrid, paucimannose and complex type, Fig 7.1) with sialylated complex *N*-glycans making up over 80% of the total glycan relative abundance.

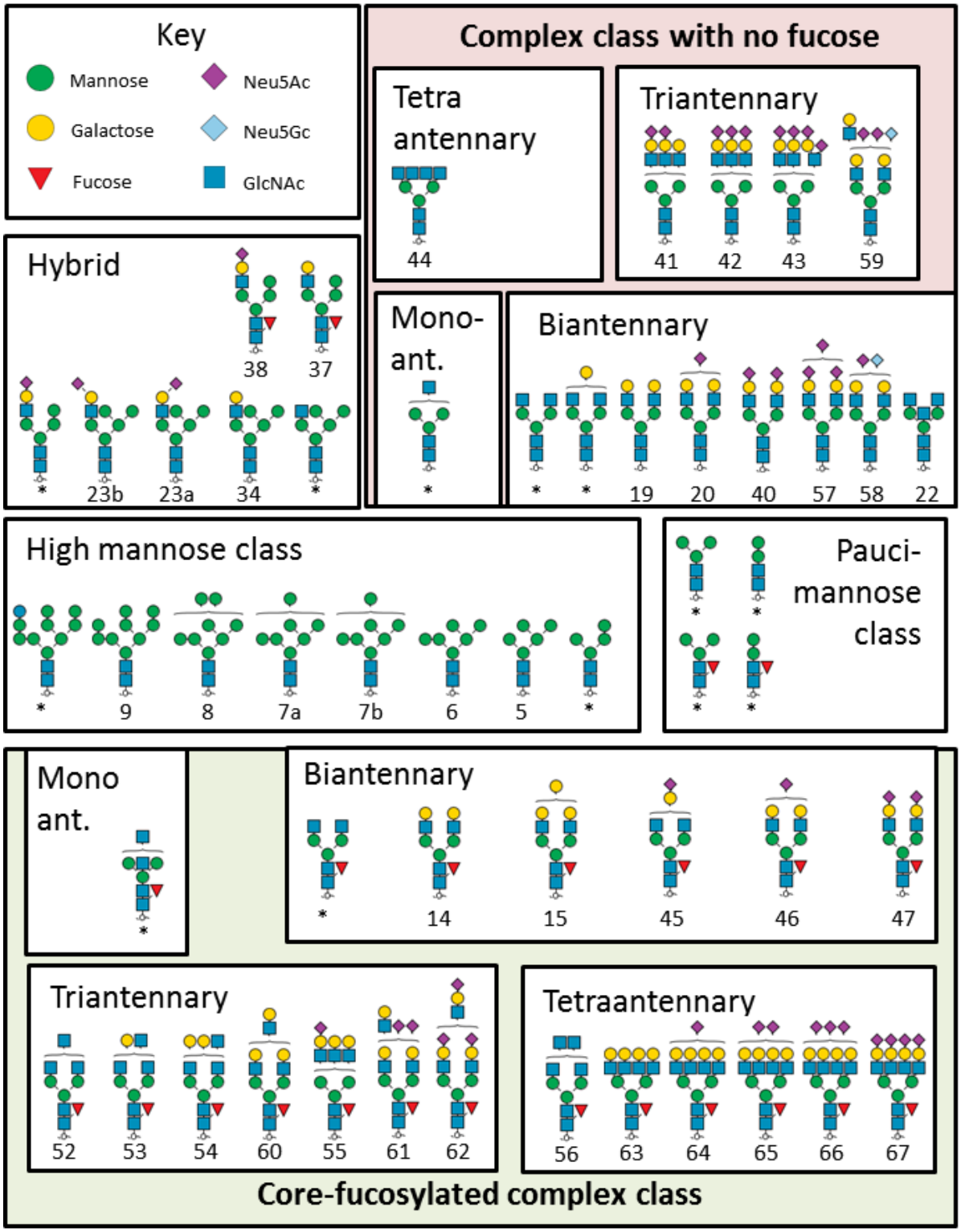


Figure 7.1 *N*-glycans detected and quantified from glycoproteins secreted by HEK293 and U87MG cell lines in serum-free media conditions, grouped by their glycan classes. * denotes glycan structures detected at trace levels without MS/MS. For this dataset, all glycans are core-fucosylated at the first GlcNAc with the fucose appearing next to the second GlcNAc due to the fucose linkage.

To put these time-course changes into perspective, we used the 0-hour LPS treatment secreted protein glycan profile as the ratio upon which a log value was determined for the relative abundance of each glycan structure. As shown in Figure 7.2, the control HEK293 cell line did not have any obvious changes of glycan abundance upon LPS stimulation, while the U87MG cell line had a clear change of *N*-glycan type abundance on these secreted proteins. Upon LPS stimulation of the U87MG

cell line, we expected glycan abundance changes in sialylation as this has been previously observed on the mucous membrane of epithelial globlet cells in a rat model[182]. Instead, we observed changes in core-fucosylation, with a 30-fold increase of glycans with this structural feature between 8 and 22 hours after LPS administration compared to the starting glycan profile. Specifically, *N*-glycan structures featuring core-fucosylation had increased expression between 10 and 22 hours after LPS treatment and this increase was complemented by an overall relative decrease in glycan structures without core-fucosylation. For the remaining time points, 33 and 44 hours after LPS treatment, the core-fucosylation reverted back almost completely to the original glycan profile, indicating this was a temporary change in the secreted protein glycosylation profile induced by LPS stimulation.

Biologically, the addition of core-fucose is mediated by fucosyltransferase 8, which has not previously been implicated in the cellular response of lipopolysaccharide treatment via the TLR4 signalling pathway. Despite this, core-fucosylation of CD14, a protein that mediates the TLR4 signalling pathway, has recently been identified to be critical for CD14-dependent TLR4 signalling[183]. As CD14 can mediate the TLR4 signalling pathway in both secreted and membrane bound forms, our results may have captured the up-regulation of core-fucosylation which, on CD14, is essential for TLR4 activation[184].

While cell lines serve as useful first avenues of investigation for human diseases, the interplay that occurs *in vivo* between different cell types and their spatial location cannot be assessed by this model system. While our *in vitro* model investigated TLR4 activation by LPS, that is a PAMP, DAMPs related to neuropathic pain can also activate TLR4[160]. Thus we moved to a mouse model of neuropathic pain to more directly investigate the effect of pain on cellular protein glycosylation.

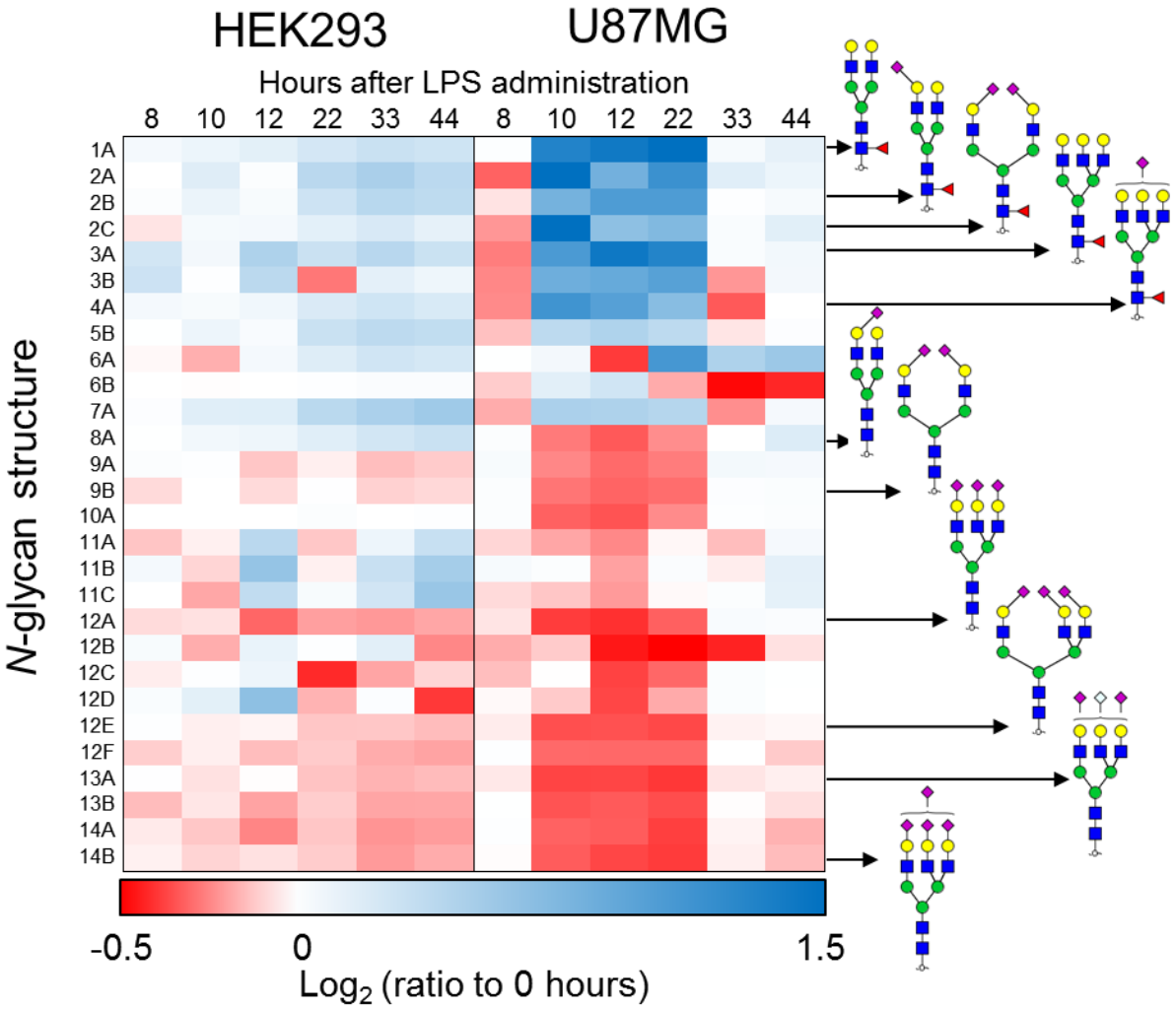


Figure 7.2 Time-course study quantifying secreted protein glycan changes following LPS administration to U87MG cells using HEK293 cells as a negative control.

7.2.2 Glycan analysis of the mouse central nervous system

Moving from the cell line model towards an animal model, we have used a mouse model to study the effect of TLR4 activation on *N*-glycan expression in both the brain and spinal cord using the CCI mouse model of neuropathic pain. These nervous system regions were chosen as they are regions involved in the perception of pain. Following peripheral nerve injury, the glia cell type can become activated in the spinal cord and cause signalling to neurons in the dorsal horn of the spinal cord which in turn transmits these signals to and from the brain[185]. The lumbar spinal cord region (L3, L4 and L5) is of particular interest for the CCI injury model as the plantar surface of the mouse hind-paw is innervated by components of the L3, L4 and L5 spinal nerves and the mouse hind-paw is used for behavioural Von-Frey measurements of allodynia and hyperalgesia (Supp Fig 7.1)[186]. The periaqueductal grey (PAG) and the rostral ventromedial medulla (RVM) transmit and modulate descending signals to the spinal cord and both of these brain regions have been implicated in the maintenance of neuropathic pain[187].

As samples were to be used as part of an interdisciplinary project that involved spatial imaging of metabolites, only a small amount of tissue was available (<5mg). Following protein extraction, using homogenisation and chloroform:methanol:water protein precipitation, the brain samples were consistent in protein amounts (Table 1) but the spinal cord samples did not have the same amount of protein, especially the L3 tissue samples which were particularly low. Samples with low glycan signal to noise ratios were excluded from data analysis with no further outlier removal.

Tissue location	**Average protein amount (ug)**
PAG	5.94
RVM	5.79
L3	1.40
L4	2.63
L5	2.75

Table 7.1 Extracted protein of representative samples from each tissue region. Protein content determined using the Bradford protein assay. The PAG and RVM locations are from the brain while L3, L4 and L5 are from the spinal cord

N-glycans released from the brain samples by PNGase F were identified and quantified by PGC LC MS/MS and found to all comprise high mannose, paucimannose, hybrid and complex glycan classes (Fig 7.3, Supp Table 7.3). The total *N*-glycan intensity, a measure of the abundance of all N-glycan species in the sample, in the periaqueductal grey (PAG) and rostral ventromedial medulla (RVM) tissue regions were approximately the same which was expected due to each sample having similar amounts of extracted protein.

The different spinal cord regions had variation in their total *N*-glycan intensities, likely caused by observed differences in extracted protein amounts, which resulted in a lower number of glycan structures detected, compared to the brain samples. Samples from the L3 tissue region had a particularly low abundance of glycans and thus could not be compared to L4 or L5 samples. Due to the variability in overall glycan intensity, tissue level comparisons were only performed between PAG and RVM brain regions as well as between L4 and L5 spinal cord regions.

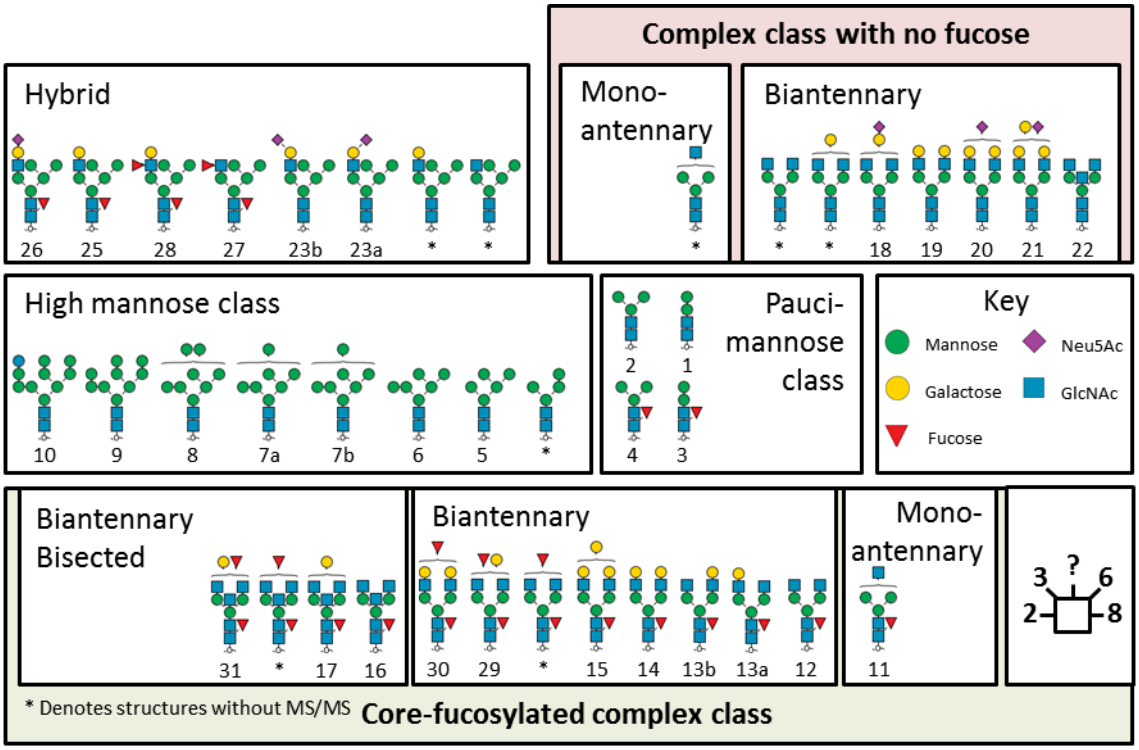


Figure 7.3 *N*-glycans detected and quantified from the PAG, RVM and spinal cord regions of mice, grouped by their glycan classes. Referred structure number denoted under *N*-glycan graphic.

7.2.3 Mouse brain tissue *N*-glycan analysis

7.2.3.1 Characterisation of N-glycans in the mouse brain

Analysing the glycans found in brain tissue regions RVM and PAG (n=11, 12 for RVM and PAG respectively. Fig 7.4 and supplementary table 7.3), a large proportion of glycans were from the high mannose class (relative abundance average across both tissue regions of 44% ±5% S.D.) and the complex class (relative abundance average across both tissue regions of 46% ±5% S.D.). Interestingly, only 3% of glycans by relative abundance were sialylated (±1% S.D.), 21% of glycans were bisected with β1,4 linked GlcNAc (±3% S.D.) and 4% of glycans were of the paucimannose class (±1% S.D.).

Due to the low sialylation and unusually high abundance of bisected *N*-glycan structures, we compared our results to previously published analyses of mouse brain *N*-glycans. We found the most comparable data set to be the Centre for Functional Glycomics dataset of a mouse whole brain. The CFG dataset is slightly different from our results as they homogenised an entire brain rather than a tissue section, thus losing brain substructure localisation (http://www.functionalglycomics.org/glycomics/publicdata/glycoprofiling-new.jsp, accessed September, 2017). Our main point of interest, for this comparison, is that the top 10 most abundant *N*-glycans found in our study of the proteins in the PAG and RVM samples are very similar to the glycans released from whole brain homogenate analysed with a different analysis platform by the CFG (Table 7.2).

With this analytical platform (Agilent 3D ion-trap), Abrahams et al observed an underestimation of sialylated glycan structures[188] and, compared to another analytical platform used in this thesis (Thermo Scientific linear ion-trap), this is a more significant issue for the 3D ion-trap architecture. Another complication is the amount of tissue used (less than 10mg for our work compared to a whole brain, as published by the CFG). We expect with an increased amount of sample, detection and quantitation of these larger sialylated glycan structures would occur.

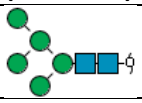
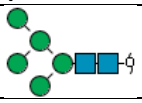
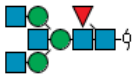
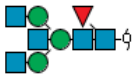
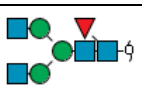
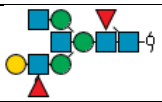
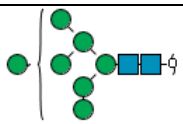
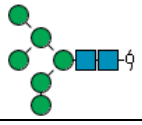
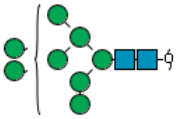
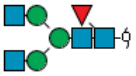
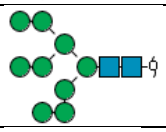
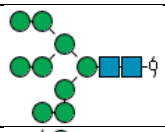
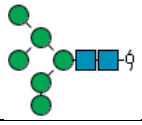
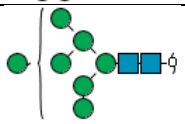
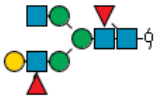
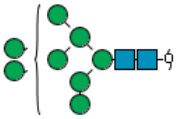
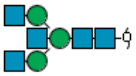
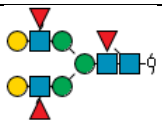
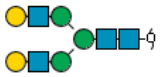
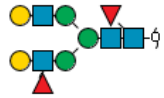
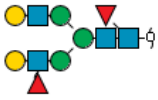
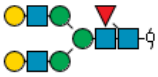
Abundance Rank	PAG *N*-glycans (Mouse brain) (This data)	Total *N*-glycans (Mouse brain) (CFG whole brain data)
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		

Table 7.2 *N*-glycan structure abundance rank comparison between this dataset and a previously published dataset of the CFG, both looking at mouse brain *N*-glycans.

7.2.3.2 Brain region-specific N-glycan differences

To see if there were any significant differences in glycan relative abundances in different regions of the brain, the PAG and RVM glycan profiles were compared.

Overall, nine of the thirty-six *N*-glycans quantified were at significantly different relative abundances between the PAG and RVM tissue regions without clustering by their treatment group (Fig. 7.4). The relative abundance of each of these glycans ranged from 1% to 8%, with glycan structures containing features such as core-fucosylation, outer-arm fucosylation and sialylation. To measure these structural features across all quantified glycans, a feature-based quantitation approach was applied.

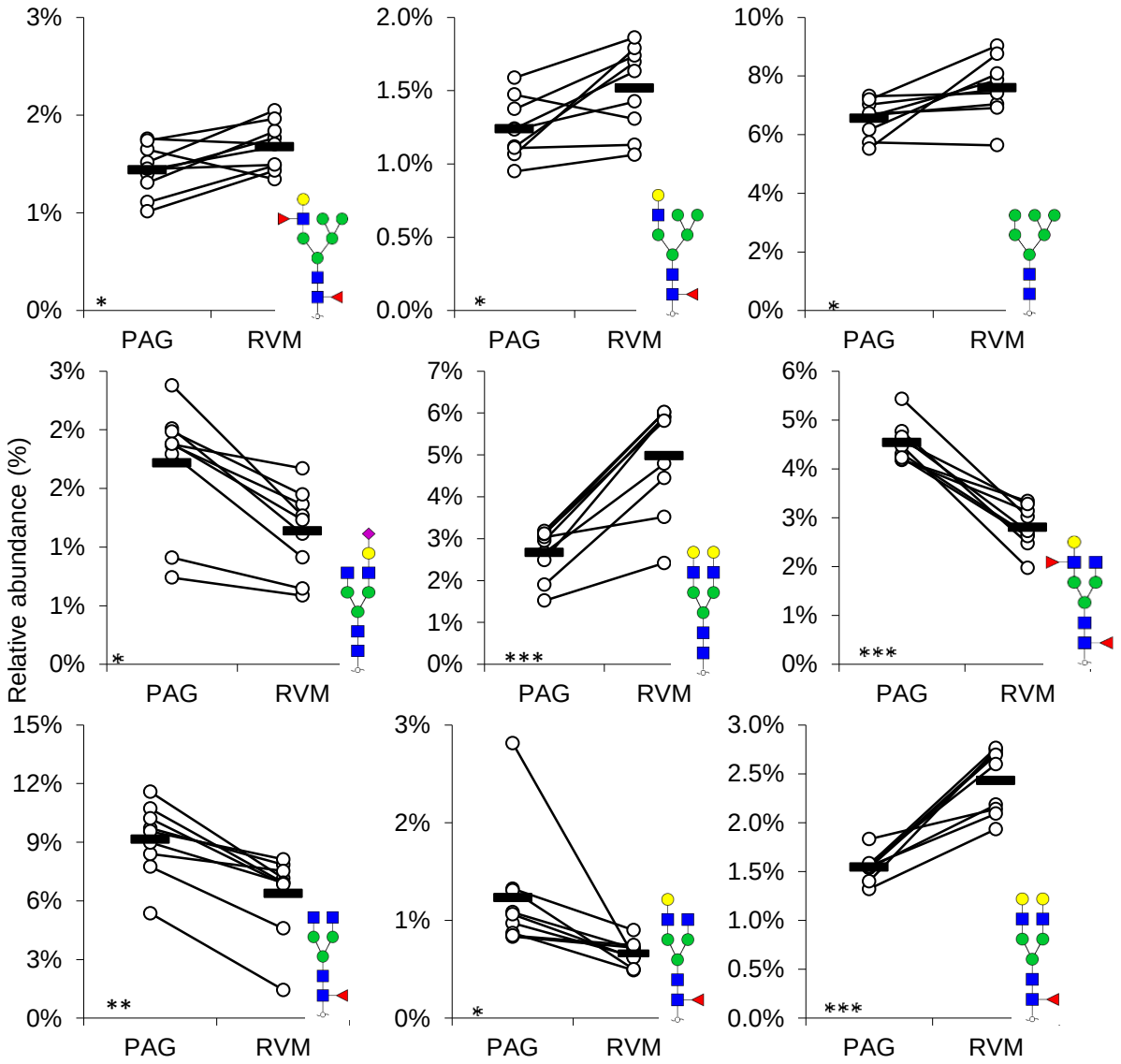


Figure 7.4 Significant *N*-glycan structural differences between PAG and RVM brain regions with tissue regions matched by individual animal. Symbols (*, **, ***) denote $p < 0.05$, 0.01 and 0.005, respectively. Each data point is an individual animal with regions from the same animal linked by a line.

Following analysis of individual *N*-glycan structures as shown in Fig 7.4, the two biggest glycosylation class modifications in our dataset were assessed: core-fucosylation and sialylation status (Fig 7.5). The PAG was found to have a greater proportion of sialylated *N*-glycans (structures 18, 20, 21, 23a/b and 26a/b) compared to the RVM while the complementary reverse was true, where RVM had a greater abundance of *N*-glycans that were the precursors for the addition of sialic acid (structures 19 and 25). The presence of core-fucosylation was also assessed but no significant differences were observed.

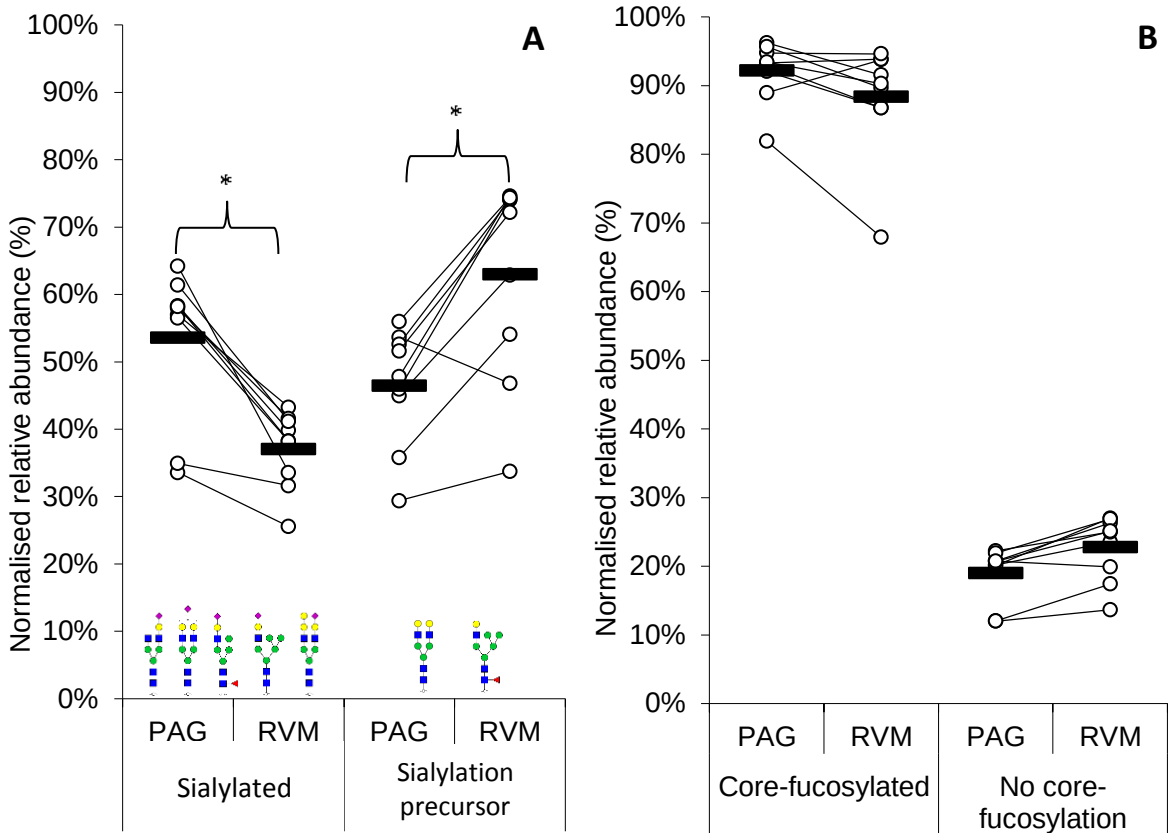


Figure 7.5 Structural glycan comparison between PAG and RVM, matched regions from the same animal. (A) Sialylation status. (B) Fucosylation status of the *N*-glycan core, Symbol (*) denotes $p < 0.05$.

When comparing glycan type class abundance (paucimannose, high mannose, hybrid and complex types) from the PAG or RVM samples between the three different treatment groups (no treatment, sham surgery and CCI surgery), no significant differences were observed. When treatment groups were excluded and all PAG samples compared to RVM samples, the paucimannose glycan type (2-5% of all *N*-glycans by relative abundance) was found to be significantly more abundant in the PAG and no other significant differences were observed for the remaining glycan classes in the two brain regions (Fig 7.6).

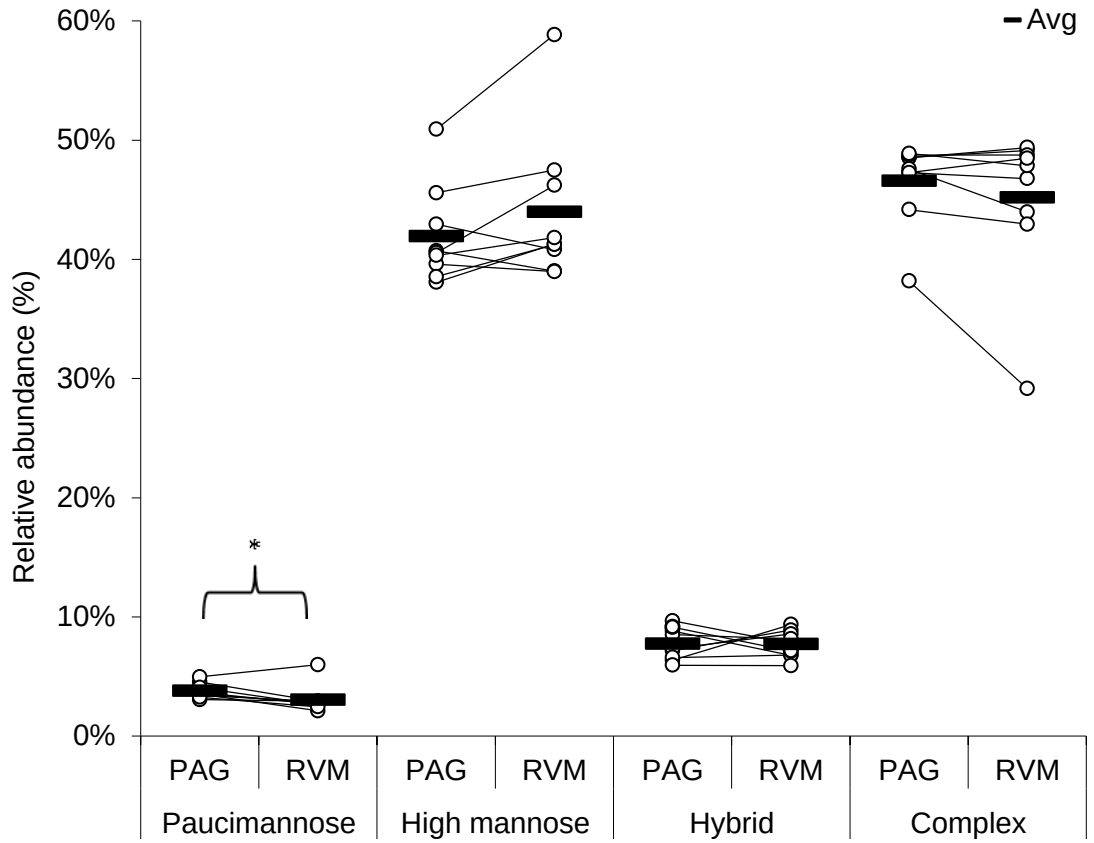


Figure 7.6 Structural glycan type comparison between PAG and RVM, matched regions from the same animal. Symbol (*) denotes $p < 0.05$.

7.2.3.3 Effect of neuropathic pain on the N-glycans in the mouse brain

To investigate the effect of induced chronic pain on the *N*-glycosylation of the brain and spinal cord proteins, three treatment groups of mice were chosen: no treatment control group, sham surgery group and CCI surgery group. The sham surgery group was required as the surgery alone will cause pain and localised inflammation, confounding the interpretation of the effects on the chronic CCI surgery group.

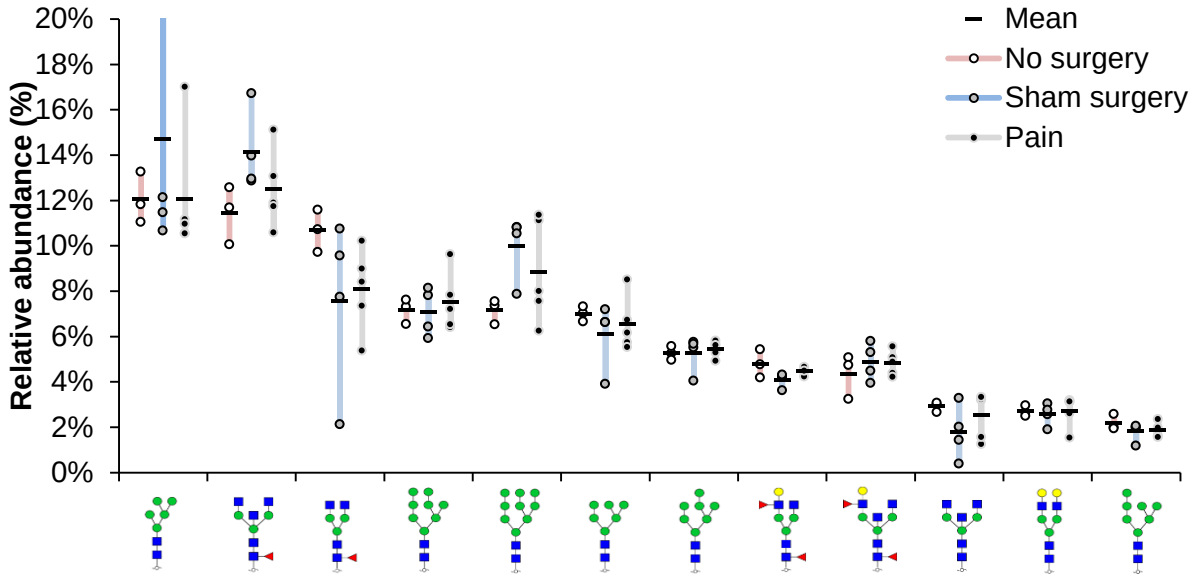


Figure 7.7 Top twelve most abundant *N*-glycans on the proteins extracted from the PAG region of the brain and their relative abundances in each replicate and treatment group

For the top twelve *N*-glycans of the PAG brain region (Full quantitative data can be found in Supp Table 7.3), one significant glycan change of the abundance of a bisected bi-antennary structure was observed when comparing sham surgery and pain treatment groups but when taken into consideration with the no surgery control, it was not significantly different (Fig 7.7). The utilisation of three different treatment groups was essential for this research as the TLR4 pathway can be activated by the surgery required for CCI injury, thus our sham surgery group was used to control for this and focus our study on chronic activation of TLR4[189]. As our aim was to understand the effect of TLR4 activation on protein glycosylation, significant glycan abundance changes between the no surgery control and CCI surgery groups complicated our findings as these changes were greater than those between CCI surgery and sham surgery groups.

Glycans from the RVM had a greater amount of variability, measured by RSD, compared to the glycans quantified from the PAG (Fig 7.7). Based on the results of the PAG and RVM glycan analysis, sham surgery has a larger statistical effect on the relative abundance of certain glycan structures (FA2B, M6, M7 and FA2G2F1) compared to the actual pain treatment. As the purpose of this study was to compare the effect of surgery with CCI injury on protein glycosylation, the surgery alone appears to cause the biggest change in protein glycosylation of the PAG/RVM brain regions rather than the chronic injury.

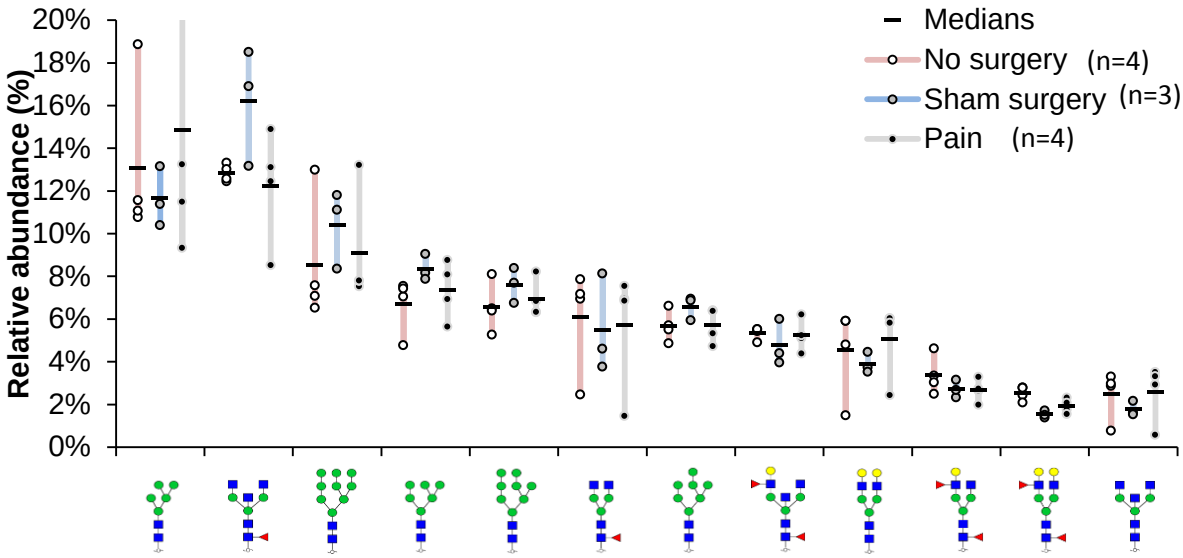


Figure 7.8 Top twelve most abundant *N*-glycans on the proteins extracted from the RVM region of the brain and their relative abundances in each replicate and treatment group

The standard deviation of the mean for each glycan assessed in both brain region samples was, on average, 25% of the mean, which made it difficult to observe any significant changes. Based on our previous work, the technical variation in peak area is estimated to be approximately 5% therefore a large component of the intra-treatment group variability is due to differences in biological factors in the individual mice.

7.2.3.4 Behavioural correlation to glycan abundance in brain tissue sections

As mice within the same treatment group had significant variation in glycan abundances, we investigated this variation by correlating glycan structure abundance with each mouse's individual behaviour in relation to pain, as measured by allodynia and hyperalgesia (Supp Fig 7.1). Evaluating the R^2 correlation value between these quantitative behavioural markers, the PAG region was found to have several glycans with significant correlation to hyperalgesia and/or allodynia (Fig 7.9). The FM2 glycan structure was significantly and negatively correlated to hyperalgesia in the PAG tissue region. Evaluating the correlation of glycan structure abundance to allodynia, FM2 ($(Fuc)_1$ $(Man)_2$ $(GlcNAc)_2$) and M5 ($(Man)_5$ $(GlcNAc)_2$) was significantly positively correlated while FA2 ($(Fuc)_1$ $(Man)_3$ $(GlcNAc)_4$) and FA2G1 ($(Fuc)_1$ $(Man)_3$ $(Gal)_1$ $(GlcNAc)_4$) were significantly negatively correlated with allodynia.

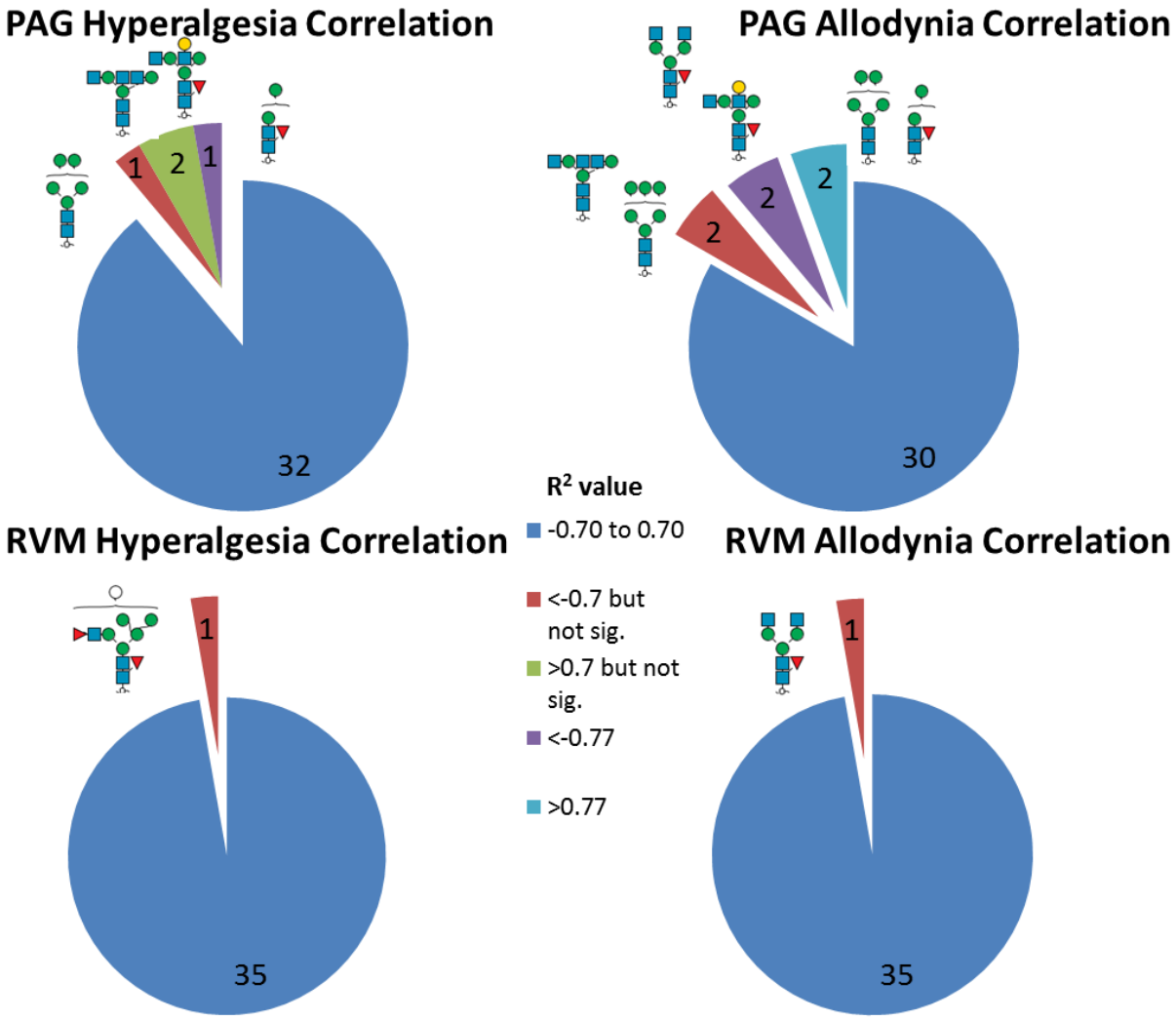


Figure 7.9 Correlation of mouse pain behaviour and glycan abundances from PAG and RVM mouse tissue sections. Values displayed are R^2 correlation values to behavioural data with n=14. Only glycan structures with abundance at least moderately correlated (<0.7/<-0.7) to behavioural values shown.

On the other hand no glycans were significantly correlated to these behavioural traits of hyperalgesia and allodynia for the RVM tissue region. This was a surprising result as both the PAG and RVM, in tandem with the spinal cord dorsal horn, constitute a neuronal circuit for sensation of pain[190].

Despite using three different experimental groups of mice for evaluating the effect of neuropathic pain on *N*-glycan quantitation on brain proteins, the biological variation within a single treatment group made the observation of specific glycan biomarkers difficult in this model of neuropathic pain. This is perhaps one reason why a less subjective pain evaluation method has not been established compared to the existing pain scale frequently used in Australian hospitals[191]. As complementary and animal specific behavioural data was collected (Supp Fig 7.1), variation in an individual animal's behaviour to CCI and its impact on *N*-glycan abundance in mouse tissue was reduced.

There was however the interesting identification of FM2, a paucimannose-type *N*-glycan, in the PAG region of the brain that was correlated to both mice hyperalgesia and allodynia. Paucimannosidic glycans have also been implicated in the function of neural progenitor cells in the sub-ventricular

zone of the mouse brain[192]. A glycan with the bisecting glycan structural feature also appeared to be correlated to the behavioural data however this was not significant for either hyperalgesia or allodynia.

7.2.4 Mouse lumbar region spinal cord *N*-glycan analysis

7.2.4.1 Region-specific differences for spinal cord tissues

At first, we evaluated treatment and lumbar region spinal cord-specific glycan abundance differences however statistical significance ($p < 0.05$) was not achieved thus we included the analysis results for all treatment groups (no treatment, sham surgery and CCI surgery) for evaluation of region-specific differences (Supp Table 7.4). The L3 tissue samples could not be included in this direct comparison due to glycan total abundance being very low in signal to noise ratio when compared to L4 or L5. L4 and L5 were compared by evaluating 15 abundant glycan structures common to the two samples. Combining the data from all treatment groups to improve statistical power, L4 had significantly increased relative abundance of glycan structure 9, and L5 had significantly increased relative abundance of glycan structure 14 (Fig. 7.10). No other significant changes were observed.

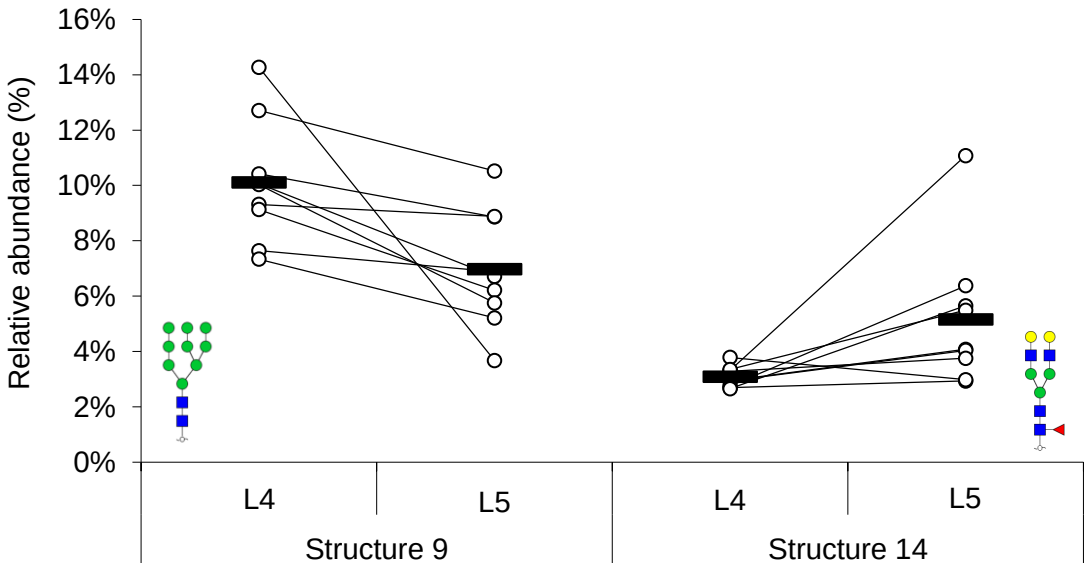


Figure 7.10 The two significantly different *N*-glycan structures between L4 and L5 with tissue regions from the same animals linked. Symbol (*) denotes $p < 0.05$.

As with the PAG and RVM samples for region-specific *N*-glycan analysis, treatment groups were pooled for improved statistical power. Evaluating the *N*-glycan types and their relative abundance for L4 and L5, L4 samples were found to have a significantly greater abundance of high mannose glycans whereas L5 samples had a significantly greater abundance of complex and hybrid glycans (Fig. 7.11). The paucimannose class was not covered in this comparison due to low signal to noise for these samples. Separating the data into treatment groups reduced the significance of these trends, with no significant differences between spinal cord regions, likely due to the comparatively low number of replicates.

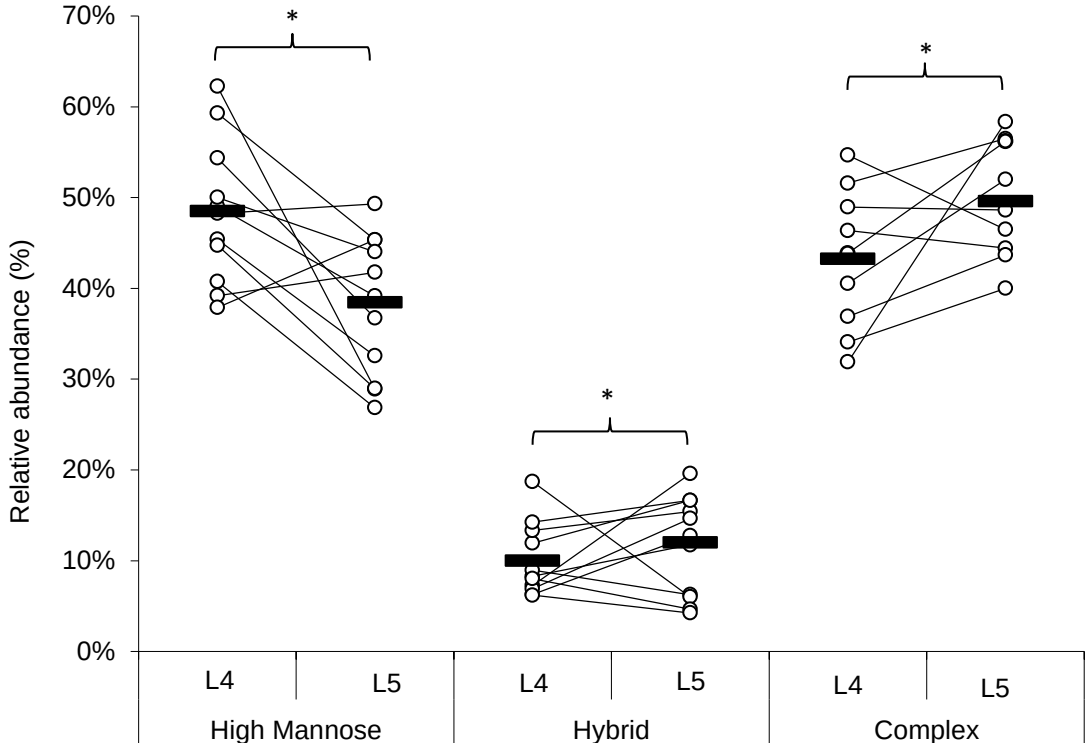


Figure 7.11 Structural glycan type comparison between L4 and L5, matched regions from the same animal. Symbol (*) denotes $p < 0.05$.

7.2.4.2 Effect of neuropathic pain on the N-glycans in the mouse spinal cord

Evaluating individual *N*-glycan abundance across the three treatment groups, no significant differences, specific for CCI treatment, were observed in L3 tissue samples (Fig 7.12). There was more variation between L3 tissue samples than all other tissue types which is likely due to the low glycan abundance. Evaluating individual *N*-glycan abundance, no significant differences were observed in L4 samples (Fig 7.13) with surgery or CCI chronic pain. In L5 samples, only one glycan was significantly different (structure 9) which increased in relative abundance in sham surgery when compared to no surgery (Fig. 7.14). No other significant changes were observed.

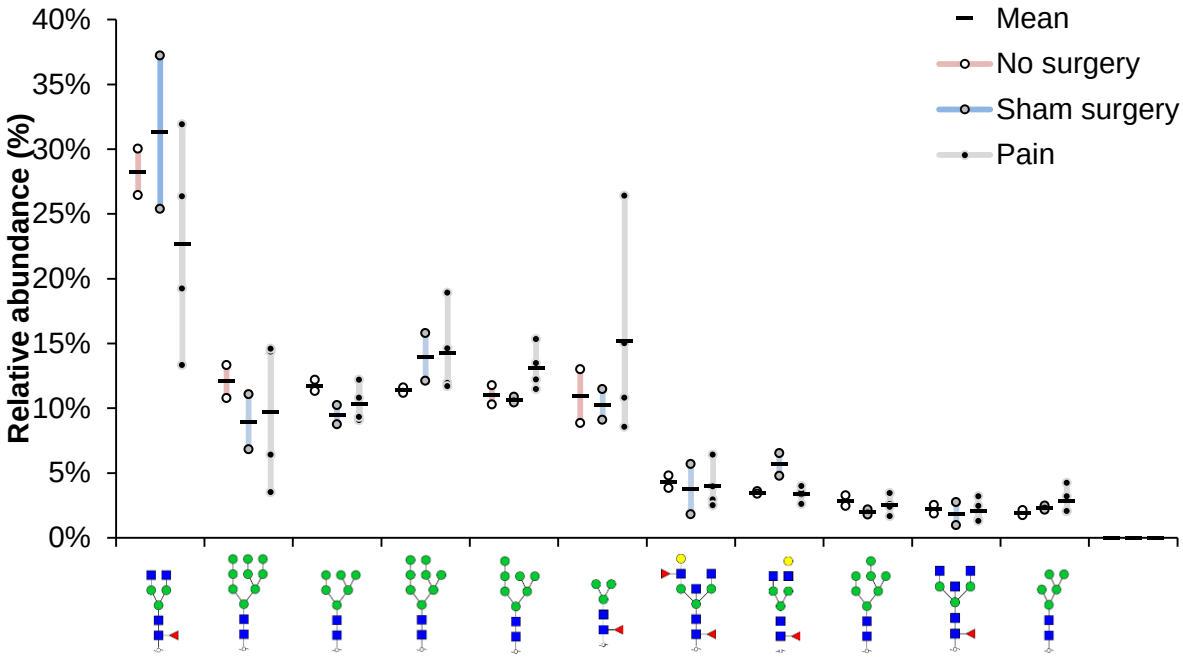


Figure 7.12 All identified *N*-glycans and their relative abundances in each replicate and treatment group on the proteins of L3 spinal cord tissue

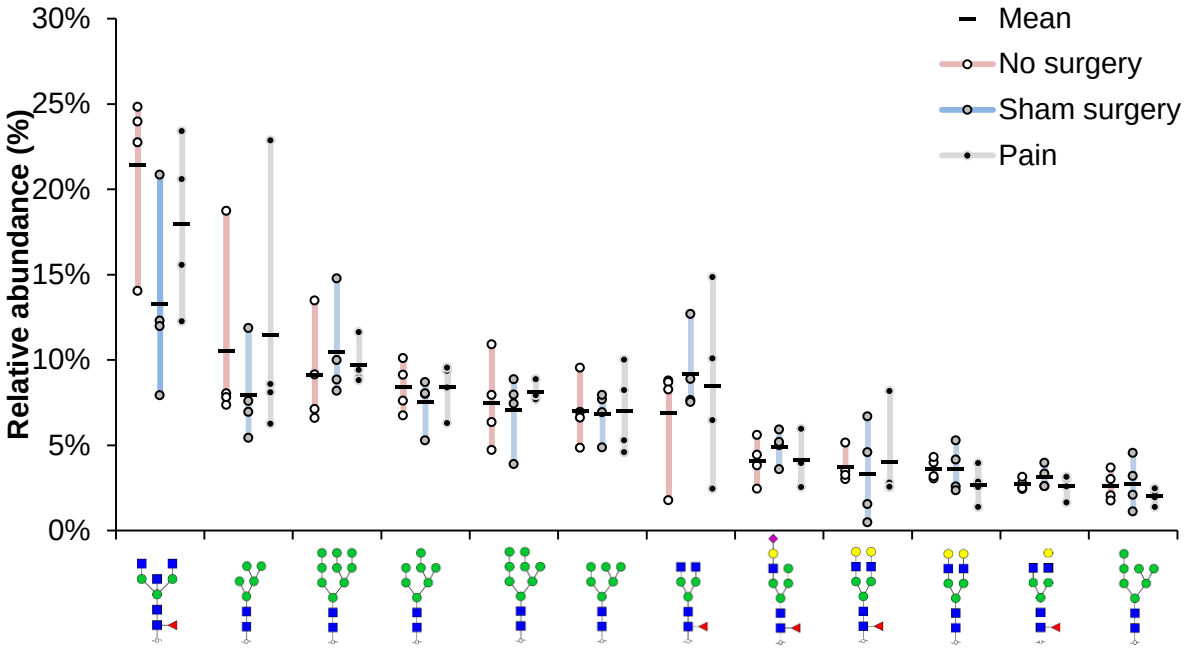


Figure 7.13 Top twelve most abundant *N*-glycans and their relative abundances in each replicate and treatment group on the proteins of L4 spinal cord tissue.

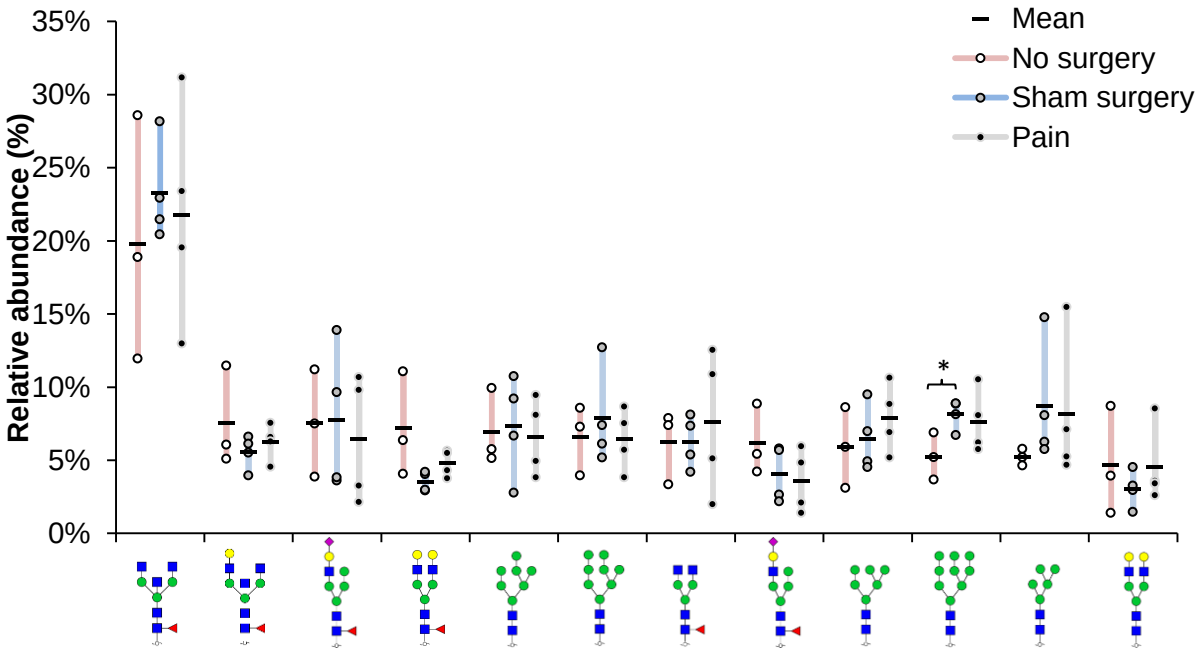


Figure 7.14 Top twelve most abundant *N*-glycans and their relative abundances in each replicate and treatment group on the proteins of L5 spinal cord tissue. Symbol (*) denotes $p < 0.05$.

7.2.4.3 N-glycans on LHS vs. RHS spinal cord tissue in response to pain

As the CCI injury is a unilateral injury to the mouse (on the right hand side of the animal), any change to protein glycosylation could either be unilateral or bilateral due to the pain signal being sent to and from the brain via the spinal cord. Evaluating opioid binding in rat spinal cords with unilateral neuropathy, Stevens *et al* found that opioid binding in specific areas of the rat lumbar spinal cord

changed upon nerve injury and these effects were bilateral[193]. This is contrasted by Honore *et al* which found that there was an increase in unilateral isolectin B4 binding to the same side as the side in which spinal nerve ligation took place (ipsilateral), which typically binds alpha-galactose residues and acts as a marker for epithelial cells[194]. We sought to evaluate if protein glycosylation changes to the spinal cord do occur and if the affect is ipsilateral, contralateral or bilateral.

To complement the existing spinal cord glycan data for different lumbar regions, a second cohort of mice were treated with the same methods previously described with all extracted tissue being used for glycan analysis only. This increased amount of material (approximately 2x) allowed glycan abundance differences between the left hand side and right hand side of the lumbar spinal cord region to be determined. Across all the samples, over 50 *N*-glycan structures were now able to be characterised and quantified as opposed to the 36 *N*-glycan structures covered in the brain regions and 15 *N*-glycan structures covered in the previous spinal cord samples (Supp Fig 7.3).

Several reports of the brain protein *N*-glycosylation repertoire have been published[195–198], but for the spinal cord, glycosylation profiling has not been published (Google Scholar and PubMed. Last checked Feb 2018). Possibly, as both the spinal cord and brain are part of the central nervous system, there is are similarities in their cell populations with glia and nerve cells abundantly present[171, 199, 200] Interestingly, sialylated species were shown to be in low abundance or not detected in the brain tissue samples (Fig 7.15) compared to in these spinal cord samples in which we observed 15 sialylated *N*-glycan structures by using a greater amount of tissue (sialylated *N*-glycans made up a total average of 8% relative abundance, Supp Table 7.5).

There were three mice per treatment group (no treatment, sham surgery and CCI surgery with a different cohort compared to the above figures) with LHS and RHS samples of the spinal cord taken for a total of 6 samples per treatment.

The ten most abundant glycans covered approximately 64% of the total *N*-glycan abundance, thus only the ten most abundant *N*-glycans were used to generate a figure comparing left hand side (LHS) and right hand side (RHS) treatment differences. The abundance differences of the remaining *N*-glycans were also assessed but no statistically significant differences were observed so were not included in the comparison. As shown in Figure 7.16, biological variation for the no treatment control group was observed to be 5-10% RSD. No significant differences were observed between LHS and RHS samples with no pain treatment.

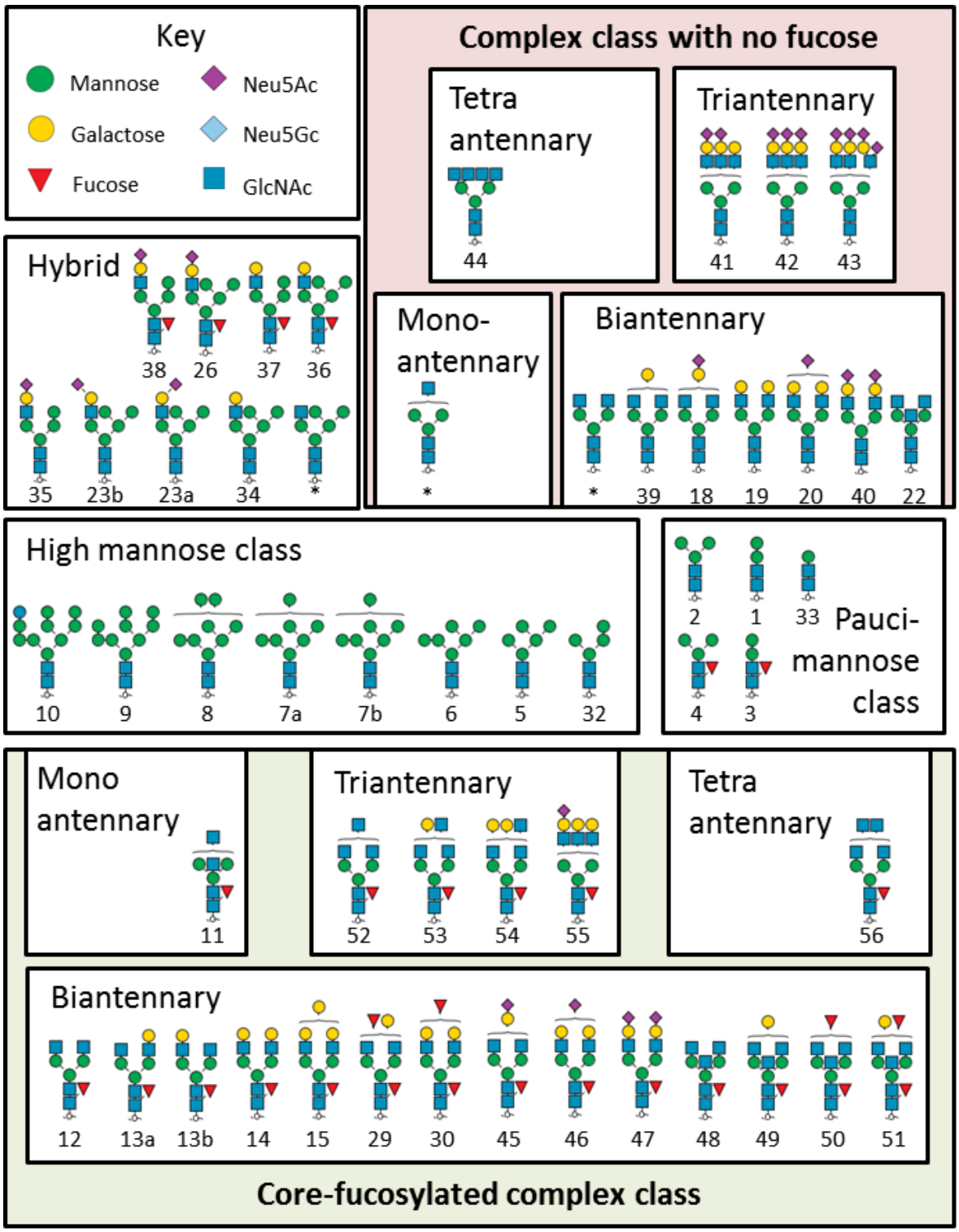


Figure 7.15 *N*-glycans detected and quantified from mouse spinal cord tissue from the lumbar region, grouped by their glycan classes. * denotes glycan structures detected at trace levels without MS/MS.

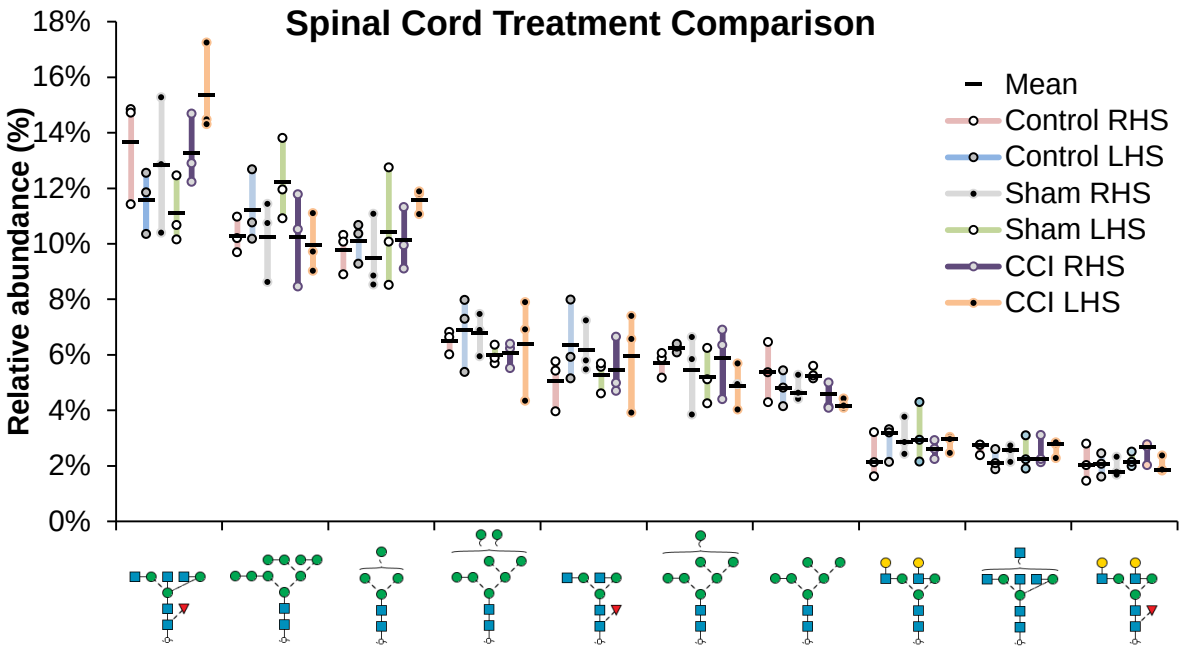


Figure 7.16 LHS vs. RHS spinal cord comparisons for top ten *N*-glycans of treatments (no treatment, sham surgery and CCI treatment)

Greater variation of glycan abundances was observed in sham surgery and CCI surgery treatment biological replicates (10-15% RSD) (Figure 7.16), which is consistent with the variation observed in the previous brain region analyses. For both the sham surgery control group and CCI surgery group, no significant differences in relative glycan abundance were observed between the LHS and RHS samples. Paired and unpaired analyses were performed with the lowest p value found to be 0.25.

However, investigating the effect of pain in relation to spinal cord side, one main significant ($p < 0.05$) difference was observed for the glycan structure FA2 (bi-antennary glycan with core-fucosylation). This structure was significantly different between LHS CCI and LHS Sham as well as between the LHS CCI and LHS control. Without seeing significant differences of more than one *N*-glycan structure between the LHS and RHS of the spinal cord between treatment groups, there is little to suggest that upon pain induction the spinal cord protein glycosylation changes based on side.

7.3 Conclusions

As a result of LPS stimulation of U87MG human glioblastoma cells, a 30-fold increase of core-fucosylated *N*-glycans occurred on secreted proteins with a corresponding decrease in non-core-fucosylated *N*-glycans. The secreted proteins were chosen as a target of profiling to identify any changes in cellular machinery without requiring cell death. This changed glycan profile appeared transient, with full recovery to normal after 33 hours after LPS-administration. As the effect of TLR4 activation on secreted protein glycosylation has not previously been reported, this serves as a useful starting point for understanding the mechanism which causes the significant changes seen in cell culture in response to LPS inflammatory stimulation.

Over thirty glycans were characterised and quantified from analysing the released *N*-glycans from proteins of central nervous system mouse tissues. The majority of the quantified glycans in the tissues were of the high mannose class which significantly differed from the cell line secreted protein glycan profiles which were largely complex glycans with several antennae.

A CCI chronic pain model was used to try to identify glycans that could be used as markers for pain states. Looking specifically at the PAG brain region, a glycan from the paucimannose class, that has been linked to inflammation, was found to significantly and positively correlate with mouse allodynia and conversely, significantly and negatively correlate with mouse hyperalgesia.

Looking at the lumbar region of the spinal cord, no significant differences in glycan abundance were observed on pain induction. With increased amounts of tissue for analysis, almost double the number of glycan structures were characterised and quantified and allowed for comparison between the different sides of the spinal cord under unilateral CCI but again no significant changes in glycan abundance were observed.

This study serves as a first glimpse of the role protein *N*-glycosylation may play in the central nervous system following TLR4 activation through stimulation with damage associated molecular patterns (DAMPs) and pathogen associated molecular patterns (PAMPs).

CHAPTER 8 - DETECTION OF THE TLR4 PROTEIN COMPLEX BY MASS SPECTROMETRY

8.0 PREFACE

In the previous chapter, we identified *N*-glycosylation changes in secreted and cellular proteins. While the cell lines are quite simple, moving towards the mouse model for study provided challenges regarding variation in animal behaviour following the same surgical intervention (CCI injury). In both cases, the TLR4 pathway is implicated as the main signalling pathway responsible for the physiological response to pain or LPS treatment. This pathway has been well characterised at the protein and ligand level, with crystal structures obtained of the TLR4 complex (TLR4/MD-2/CD14) with and without identified inhibitors of the TLR4 pathway.

The start of the TLR4 signalling pathway, the TLR4 complex serves as a starting point to quantify changes in the abundance of these proteins and determine how the abundance of TLR4/MD-2 is impacted by the cellular response to TLR4 activation. Previous approaches to this question utilise antibody-based assays to detect and quantify these proteins however these assays rely on assumptions regarding antibody specificity, which may not be true for the sample analysed.

Overall, we aimed to detect TLR4 and MD-2 abundance in model cell lines of TLR4 pathway activation using widely available proteomics methods of LC-MS/MS. CD14 was excluded from these assays as LC-MS-based methods for its detection have been previously developed[201].

8.1 INTRODUCTION

Proteomic approaches to protein quantitation

One of the most ubiquitous techniques for direct protein detection and identification is mass-spectrometry. Due to current technological limitations, detection and quantification of intact proteins in complex mixtures is difficult, thereby requiring the sample preparation step of digestion into peptides. This strategy, known as bottom-up proteomics (covered in section 1.1), has inherent limitations owing to the loss of proteoform information which is an issue for post-translational modifications such as glycosylation[11]. From protein digestion and analysis of the resultant peptides, peptide sequences can be matched to their precursor protein if the genome is known[202]. Mass-spectrometry has the benefit of built-in validation using MS/MS spectra to confirm peptide sequence in combination with observed precursor peptide mass[203]. The abundance of detected proteins can then be relatively or absolutely quantified using label-free spectral counting or standard dilution curves, respectively[204]. To reduce technical variation and increase throughput, several labelling techniques have been established (such as iTRAQ[205], TMT[206] and SILAC[207]) which

allow multiplexing of samples and subsequent quantitation, with TMT mass tags providing quantitation based on reporter ions[208].

LC-MS experiments are frequently used in the field of proteomics for peptide detection, quantitation and characterisation. The popularity of this setup is due to the LC, in combination with a suitable stationary phase, being capable of separating complex peptide mixtures and the MS being capable of detecting and characterising (mainly tryptic used) peptides within a complex mixture[209]. With the advent of modern computer control of mass-spectrometers, the scheduling of detection (MS1), quantitation (MS1[210], MS2[211] and MS3[212]), as well as characterisation (MS1[213], MS2[214] and MSn[215]) can be carefully controlled to answer the hypothesis that is to be tested, or generate a hypothesis. Currently, there are two prevailing methods for non-targeted peptide analysis: data dependent (aka IDA or shotgun), which ideally fragments individual peptides based on criteria such as intensity[216], or data independent, which fragments multiple peptides within a wider *m/z* range than typically used for data-dependent experiments (> 15 *m/z*)[217].

Targeted proteomics to test hypotheses

Targeted proteomics exists as an alternative to untargeted proteomics, requiring a hypothesis before data acquisition by generating a target list of precursor and/or product ions. The requirement of a target list, while a disadvantage for discovery experiments, provides improved dynamic range and accuracy for quantitation, and also continuous acquisition of MS/MS spectra for *m/z* ranges representing the target peptides[65].

The development of this list of peptide targets can be informed by results of prior untargeted analyses (self-generated or published proteomics datasets), knowledge of the protein network of interest and literature research[218]. Furthermore, practical instrument considerations can be taken account, identifying the maximum number of targets for a limited instrument duty cycle, which must be carefully adjusted[65]. Thus, the initial target list often requires classification and optimisation based on how much value each peptide ion target can provide towards answering the hypothesis[219].

One strategy for target classification in a targeted proteomics assay is the use of enzymatically digested recombinant proteins, representing the network of interest, or synthetic peptides, representing the protein sequences that are part of the target network. Analysing these simple peptide mixtures using discovery-based proteomics identifies peptides best suited towards the LC-MS platform used, giving the highest signal with a peptide sequence specific for the target protein. Another area of interest is the identification of quantotypic peptides to represent a protein, that is, the level of the measured peptide is stoichiometric to the level of the protein[220]. Ideally, this

would be a peptide that is unmodified in both sample preparation (no cysteines or methionines) and biology (no possible post-translational modifications such as glycosylation, phosphorylation and deamidation etc)[219]. As a result of target classification, a list of peptide targets that best represent the proteins of interest for a quantitative assay can be refined.

For targeted proteomics, the advent of high resolution and accurate mass (HRAM) analysers has led to the development of alternative targeted mass-spectrometry methods. There are three types of targeted proteomics approaches: SRM (targeted precursor ions and targeted product ions)[219], PRM (targeted precursor ions and untargeted product ions)[65] and DIA (untargeted precursor ions and untargeted product ions, with targeted data analysis)[217]. While SRM is associated with the triple quadrupole mass-spectrometer[219], both PRM and DIA can be performed on all HRAM mass spectrometers although are labelled with brand-specific marketing terms (MRM-HR and PRM are fundamentally the same[221], as is SWATH and Mass Spectrometry Eternal (MSE)[58, 217]).

Historically, targeted proteomics has been frequently used as SRM assays for quantifying proteins in a network of biosynthetic pathways. A study attempting to establish a mathematical model of the citric acid pathway in yeast is one notable example of SRM use, due to the requirement for absolute quantitation of enzymes in this pathway, when grown in different carbon sources[222]. They reported detection and quantitation of cell cycle proteins that were of low abundance and almost invisible to un-targeted proteomic methods for quantitation[222]. Labs that typically use SRM have begun moving towards PRM with studies demonstrating PRM achieves similar performance metrics (accuracy, dynamic range, sensitivity and reproducibility), as the untargeted product ion analysis allows reinterrogation of data to identify the most informative product ions to target, post-acquisition[223].

A relevant example for us is a study using label-free PRM to study several signalling pathways in mouse microglia in response to LPS[224]. Although the authors did not cover the toll-like receptor signalling pathway, they quantified protein changes in other inflammatory and metabolism pathways such as proinflammatory cytokine production and mitochondrial function. Using scheduled precursor selection with PRM, they were able to monitor over 450 peptides in a single run[224]. This demonstrates that PRM can be a suitable tool as a reproducible and quantitative assay to study protein networks as part of biological pathways.

As mentioned in the preface, the typical method to quantitate changes in the TLR4 pathway is an antibody-based approach. This method, immunohistochemistry, utilises anti-TLR4 and anti-MD-2 antibodies and has been the predominant method since 2000. To date, there have been no publications specifically studying TLR4-complex protein abundances using mass-spectrometry (PubMed and Google Scholar searches, Feb 2018), despite the complementary value that mass-

spectrometry can provide towards understanding the role of TLR4-complex protein abundances in the human immune response.

Harvesting data from proteomics repositories

Proteomics is one of the most active research areas in life sciences with continual advancement in instruments and techniques to identify and characterise growing numbers of proteins[225]. Accompanying this active research area is the availability, and sometimes required utilisation, of repositories for proteomic data from scientific publications. PRIDE is one example of a proteomics data repository which contains the protein/peptide identifications, supporting mass spectra evidence and metadata surrounding the published proteomic study, including experimental methods[226].

The availability of such a database allows researchers to search for proteins that are of interest and review the accompanying experimental methods used for their detection. The availability of raw data for subsequent re-analysis provides an opportunity for validation of reported, and discovery of unreported, protein identifications. The re-analysis of this data is likely the most significant bottleneck with using a repository such as PRIDE as it is often in a vendor-specific format. Several tools are suitable for the re-analysis of vendor-specific data such as PRIDE Inspector, a universal visualiser tool for proteomics data[227], and Skyline, a document editor for analyser MS-based experiments[109]. While PRIDE Inspector has the benefit of being developed as part of the PRIDE ecosystem and therefore is centred on the archive itself, Skyline is developed more generally around mass-spectra generated by proteomics experiments.

Previous proteomic characterisation of TLR4 and MD-2

The first TLR4-complex crystal structures were described by Kim *et al* in 2007, uncovering the structure of mouse MD-2 bound to mouse TLR4 with a bound small molecule antagonist known as Eritoran [157]. They identified that TLR4 interacts with MD-2 through a strictly conserved patch and Eritoran binds to the hydrophobic pocket of MD-2, with no hydrogen bonding shared between it and TLR4[157].

For the human proteins, Park *et al* studied the crystal structure of human TL4 and human MD-2 with bound LPS from *E. coli*[228]. Interestingly, only LPS resulted in dimerization of TLR4 whereas Eritoran did not, highlighting the difference between their mode of actions as an agonist and antagonist, respectively[157, 228]. Another aspect of the aforementioned crystal structures by Park *et al*, is the expression host. They utilised a Hi5 insect cell line for expression of these recombinant human TLR4 and MD-2 so that several *N*-glycosylation sites (TLR4: N35, N173, N205, N497, N526, N575, N624,

N630, MD-2: N26, N114) had paucimannose or further truncated *N*-glycan structures observed in the crystal structures[228]. Their use of a proteins with insect glycosylation therefore limits the interpretation of the glycans likely to be observed in a mammalian system[229]. These findings confirmed MD-2 as a binding partner of TLR4 and also that LPS is bound within MD-2 for subsequent TLR4 activation and dimerization.

Da Silva Correia *et al* took this one step further, experimentally examining the function of the *N*-glycosylation sites of TLR4 and MD-2 using site-directed mutagenesis and enzymatic glycan release on proteins expressed in HEK293 and HeLa cells[230]. For MD-2, they found the presence of *N*-glycosylation sites had no effect on secretion whereas TLR4 required N526 or N575 for transport of the protein to the cell surface. Functionally, the loss of MD-2 glycosylation sites resulted in suppression of LPS-induced activation of IL-8, a cytokine typically involved in the TLR4 pathway. The same outcome was observed in this study with TLR4 mutants lacking *N*-glycosylation sites (Mutant 1.7 N526A, Mutant 1.8 N575A, Mutant 2.4 N497A/N575A, Mutant 3.3 N497A/ N526A/N575A, Mutant 4.2 N497A/N526A/N575A/N624A) with N526 and N575 identified as essential *N*-glycosylation sites for cell surface expression[230].

FLAG-tagged proteins and antibodies specific for the FLAG tag have been used for antibody-based detection of TLR4/MD-2 and this technique has been an indirect method for quantifying changes in the TLR4 pathway[230].

Frequently used assays for TLR4 pathway characterisation

The ability to detect and quantify secreted and membrane-bound proteins involved in the start of the TLR4 pathway is important for studying its role in the human immune response. The most common technique is immunohistochemistry involving anti-TLR4 and anti-MD-2 antibodies. These techniques were first utilised in 2000, uncovering the LPS endotoxin response and the hypo-responsiveness that can occur as a result of randomly occurring strain-specific mutations of the TLR4 gene[231–233]. Since then, anti-TLR4 antibodies have been used extensively to identify agonists of TLR4[234–237], characterise the spatial distribution of TLR4[238–241] and understand the physiological effects of TLR4 stimulation[154, 242–244].

A recent concern regarding the use of antibodies for life science research is the validation of their specificity, reproducibility and sensitivity[245]. In our evaluation of the experimental methods of the studies cited earlier using anti-TLR4 antibodies, only 2 out of the 12 studies presented data showed specificity testing for the anti-TLR4 antibody used and surprisingly, 2 of the studies did not identify the source of the antibody (commercial or otherwise) or what the antibody was raised against. Furthermore, no validation was performed with an orthogonal technique, such as mass-

spectrometry, to confirm the detection of TLR4 instead of a protein with a similar molecular weight[246]. Despite this, these antibodies have proven useful for understanding the function of TLR4 and have typically been complemented by qualitative and quantitative mRNA assays.

Another frequently used method for assessing TLR4-pathway activation is the use of reporter cell lines which induce the secretion of reporter proteins following NF-κB activation. This technique can be performed using commercially available cell lines and allows real-time detection of NF-KB activation which has proven useful for studying TLR4 agonists[247, 248, 257–259, 249–256] and the physiological effects of TLR4 stimulation[260–264]. This technique can be seen as complementary to antibody-based TLR4 detection as it measures downstream NF-KB activation, which can however also be achieved through activation of other toll-like receptors[265].

Both antibody detection and reporter cell lines involve indirect detection of TLR4. For antibodies, it is detection by a possibly non-specific anti-TLR4 antibody, and for the use of reporter cell lines, it is the secretion of the reporter gene linked to downstream NF-KB activation. While these techniques are frequently used due to their practical value, direct detection methods that can quantify the amount of TLR4, and the partners in its complex, are of immense scientific value due to the ability to quantify the protein complex identified to be at the forefront of the immune response towards DAMPs and PAMPs.

8.2 Results and Discussion

Methods for the results generated in this section can be found described in detail in Chapter 2.

The commercially available HEK293 +TLR4 +MD-2 +CD14 (stably transfected with these genes) cell line is frequently used for studying the TLR4 signalling pathway and we have used a non-transfected HEK293 cell line as a control[266–272]. The plasmids containing the TLR4 gene confers Blasticidin resistance and the MD-2 and CD14 gene confers Hygromycin resistance, ensuring cells cultured in media containing these antibiotics possess the TLR4, CD14 and MD-2 genes. Typically, measuring perturbations of the TLR4 pathway is based on measuring the downstream response by NF-KB activity. NF-KB is not specific for the TLR4 pathway and is intracellular, making it not an ideal measurement target when evaluating changes to the early-TLR4 pathway. Thus we aimed to directly detect and quantify human TLR4 and MD-2 proteins directly using mass-spectrometry.

8.2.1 Detection of TLR4 and MD-2 from model cell lines used in pain research

In attempts to detect TLR4 and MD-2, two protein extraction approaches were initially taken. The first, whole cell lysis and cutting out regions from SDS-PAGE separations at the mass regions that TLR4 and MD-2 are expected (90-160kDa, 15-30kDa, respectively), and the second, subcellular fractionation with the whole gel lane excised after subsequent SDS-PAGE (Methods described in more detail subsection 2.4, Fig 8.1). All gel containing protein samples were then digested with trypsin and analysed with C18-uHPLC-ESI-MS/MS.

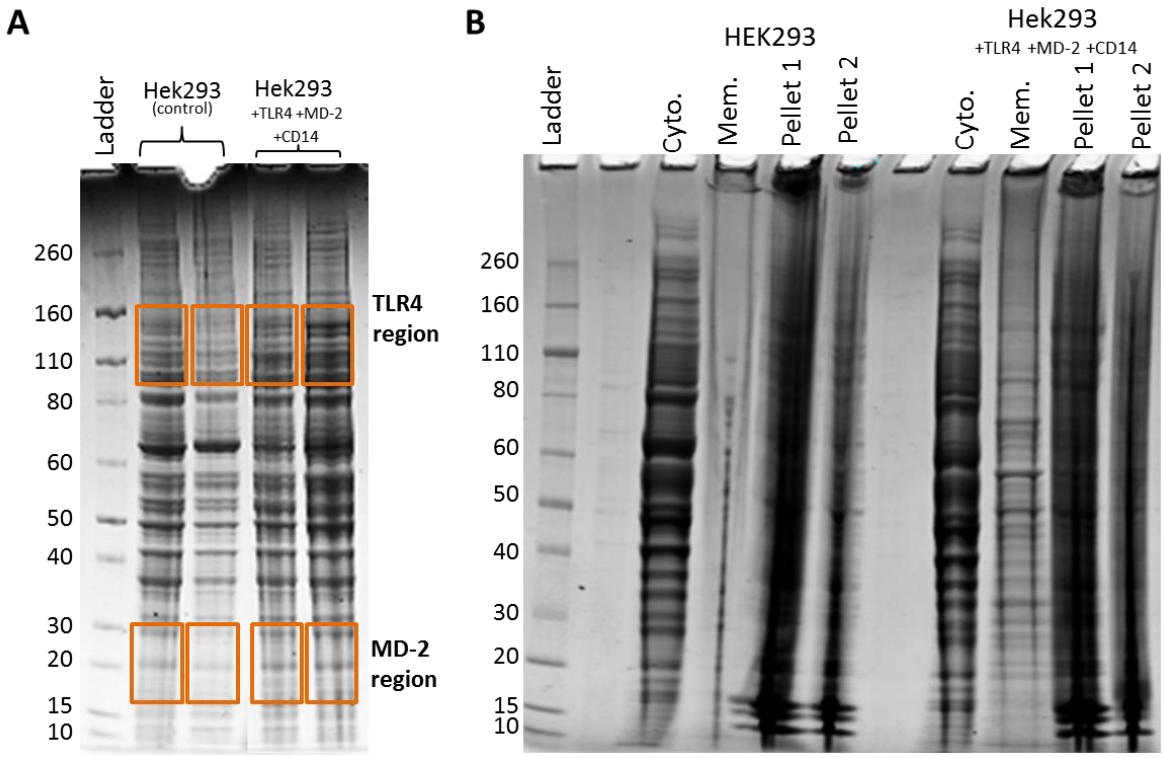


Figure 8.1 SDS-PAGE of extracted proteins from HEK293 and stably transfected HEK293 cell lines. A) Cell lysate with boxes highlighting regions cut out for subsequent mass spectrometric analysis for TLR4 and MD-2. B) Subcellular protein fractions of both HEK293 cell lines.

Comparing the results of the subsequent shotgun proteomic analysis, subcellular fractionation into cytosolic, membrane and, cytoskeletal and nuclear fractions achieved reduction of sample complexity with approximately 1500 unique protein identifications per cell line (Fig 8.2A). In attempts to reduce sample complexity, gel bands corresponding to TLR4 and MD-2 were cut out but there were no additional unique protein identifications (data not shown). From the subcellular fractions, only 2 proteins of the TLR4 signalling pathway were identified, namely CD14 and MAP kinase kinase 7 (MKK7) (Fig 8.2B). Proteins at the start of the signalling pathway, TLR4 and MD-2, were not detected in any of the samples (including the cut-out gel bands consistent with the molecular weights of MD-2 and TLR4).

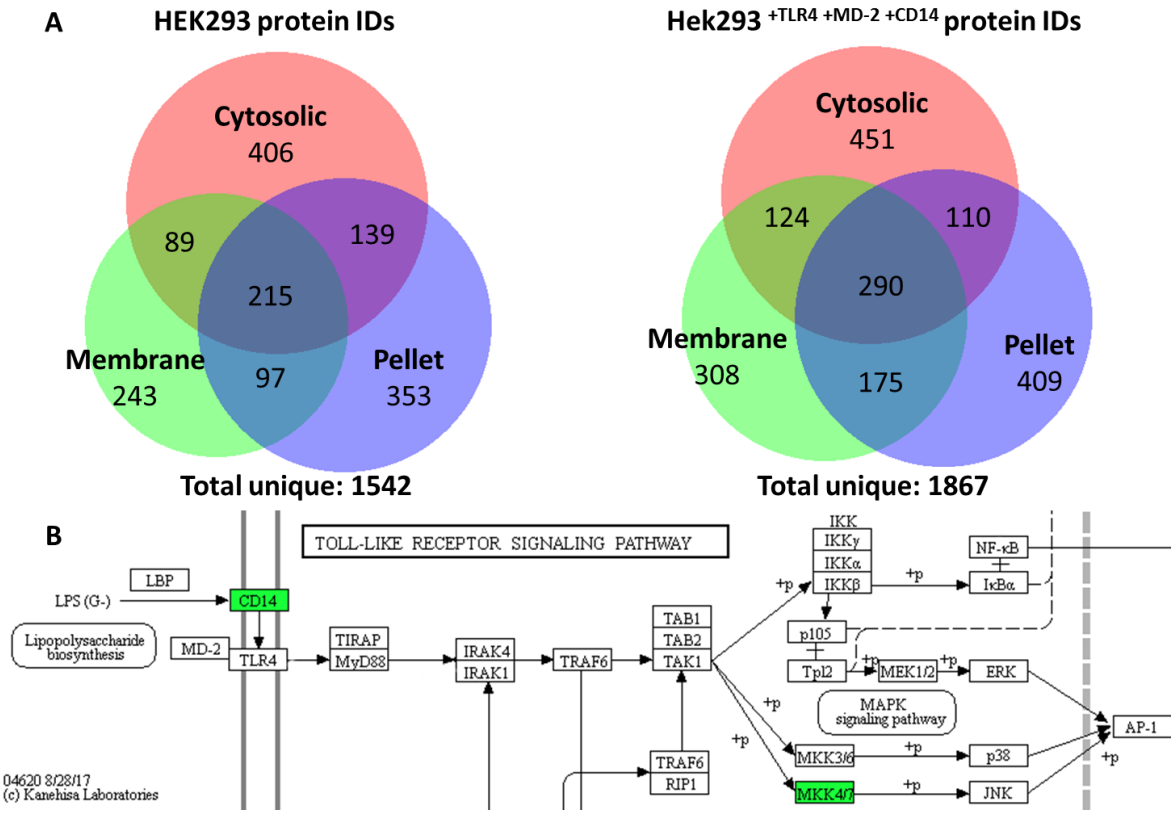


Figure 8.2 Shotgun protein identifications of subcellular fractions from HEK293 cell lines. A) Comparison of subcellular fractionation for increasing unique protein identification compared to unfractionated samples. B) Protein components of TLR4 signalling pathway with detected proteins highlighted in green

Concerned that our stably transfected cell line did not transcribe the TLR4 or MD-2 genes, although they were expected to be abundant in such a stably transfected cell line, we performed reverse transcription polymerase chain reaction (RT-PCR) to confirm the presence of mRNA encoding human TLR4 and human MD-2 in these cells. Using primers based on the plasmid sequences from the cell line vendor, we confirmed human TLR4 and MD-2 were being transcribed by the stably transfected cell line (Fig 8.3). Observed PCR products were consistent with their expected sizes, 372bp for human TLR4 and 556bp for human MD-2.

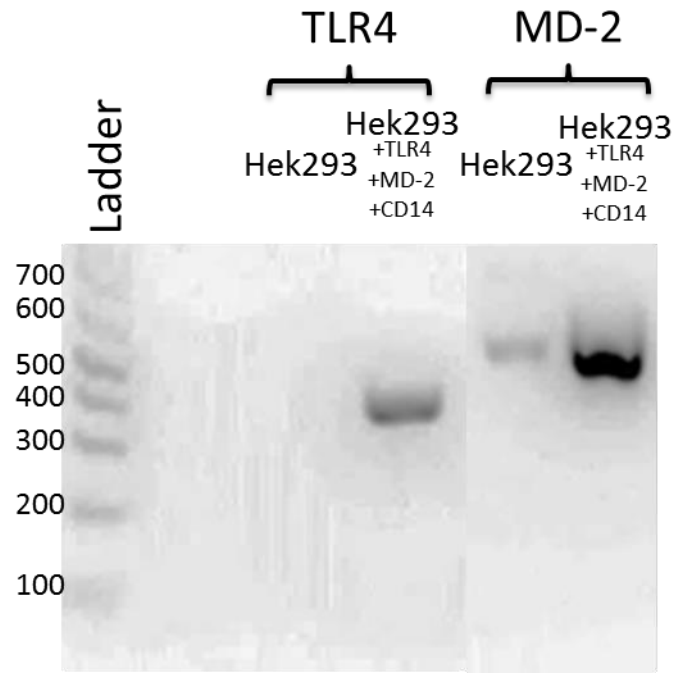


Figure 8.3 Confirmation of human TLR4/MD-2 transcription in the stably transfected HEK293 cell line using RT-PCR

Human TLR4 was found to not be transcribed in the non-transfected control HEK293 cell line and only a faint band appeared for MD-2 (Fig. 8.3), confirming the HEK293 cell line to be a suitable control due to lack of TLR4 mRNA. These findings are consistent with prior publications regarding TLR4 mRNA in HEK293 cells[273]. We thus validated that the HEK293 cell line was stably transfected with TLR4/MD-2genes and transcribed these genes to mRNA in preparation for translation to protein.

To optimise obtaining mass-spectrometric evidence of these proteins we then characterised recombinant TLR4 and MD-2.

8.2.2 Characterising recombinant TLR4 and MD-2 to inform our detection strategy

To overcome the negative results of TLR4 detection using untargeted proteomics, we purchased recombinant (HEK293 expressed) human TLR4 and MD-2 to optimise the parameters for targeted proteomics strategies of detection. Human sequence MD-2 expressed in *E. coli* was also purchased as this was devoid of protein glycosylation and much cheaper for initial instrument optimisation compared to MD-2 expressed in HEK293 cells. As shown in Figure 8.4, the purchased HEK293 proteins were found to not be of high purity, despite the vendor description of a purity >80%. To confirm the bands highlighted in Fig8.4A were MD-2 and TLR4, the bands were excised, digested to peptides and analysed using C18-LC-ESI-MS/MS. Bands 1, 2 and 3 were all confirmed to contain

human MD-2 while band 5 was confirmed to contain human TLR4. Band 4 was possibly another glycoform of MD-2 however MD-2 was not the most abundant protein in the analysed band 4 and had lower protein coverage compared to bands 2 and 3, thus it was not likely to be MD-2 (Supp Table 8.1).

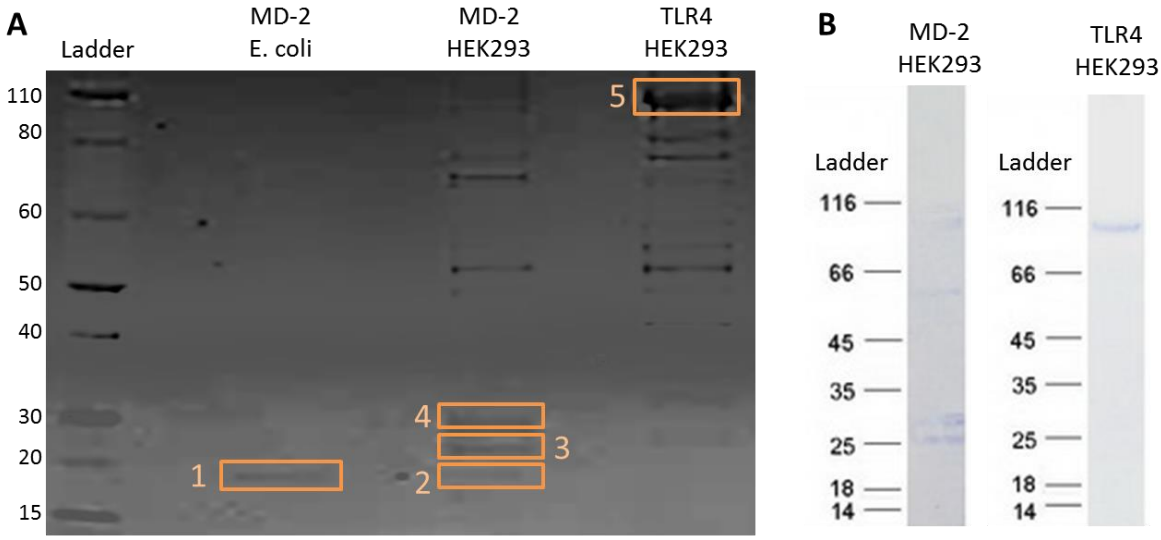


Figure 8.4 SDS-PAGE analyses of recombinant proteins, MD-2 and TLR4. A) Separation of purchased recombinant proteins from *E. coli* and human HEK293 cells with highlighted regions for subsequent mass spectrometric analysis. B) Manufacturer supplied SDS-PAGE images of purchased recombinant protein.

Looking in more detail at the TLR4 and MD-2 (recombinant protein) proteomics results, both proteins were detected with >40% peptide coverage (Table 1). Observed tryptic peptides included those with sequences featuring an *N*-glycosylation motif, but no glycosylation was initially determined due to our search method not utilising software capable of glycopeptide identification (Proteome Discover 2.2 used without Byonic).

Protein (HEK293 derived)	Band ID	Unique peptides	Modified peptides	Coverage	AA range	*N*-glycosylation sites covered
Human TLR4	5	110	68	52%	434-839	2/2
Human MD-2	4, 3, 2	19	13	48%	77-160	3/9

Table 8.1 Protein coverage statistics from mass-spectrometric analysis of peptides from digestion of HEK293-derived recombinant protein bands highlighted in Fig 8.4.

Based on the detection of non-glycosylated peptides with *N*-glycosylation motifs, and the reporting of MD-2 and TLR4 crystal structures with *N*-glycosylation[157, 228, 274], we sought to identify glycan structures attached to these proteins. Unfortunately, no released glycans could be detected from an *in-gel N*-glycan release of HEK293 expressed recombinant TLR4.

Two glycan lists were then used with Byonic proteomics software for glycopeptide-oriented data analysis, a high-mannose only list based on the glycomics results (Fig 8.6D) for recombinant HEK293 expressed MD-2, for quicker searches, and a total mammalian list, to ensure we were not missing glycopeptides containing glycans not detected with our glycomics approach. Both human MD-2 and TLR4 were observed to contain high mannose *N*-glycosylation on their peptides with no complex *N*-glycans detected on either protein (Fig 8.5). For MD-2, all previously found *N*-glycosylation sites were found to be occupied by high mannose *N*-glycan structures. For TLR4, glycopeptides observed only covered 3 or 4 of the possible *N*-glycosylation sites with some uncertainty for the peptide QGMPVLSLNITCQMN which contains two possible *N*-glycosylation sites.

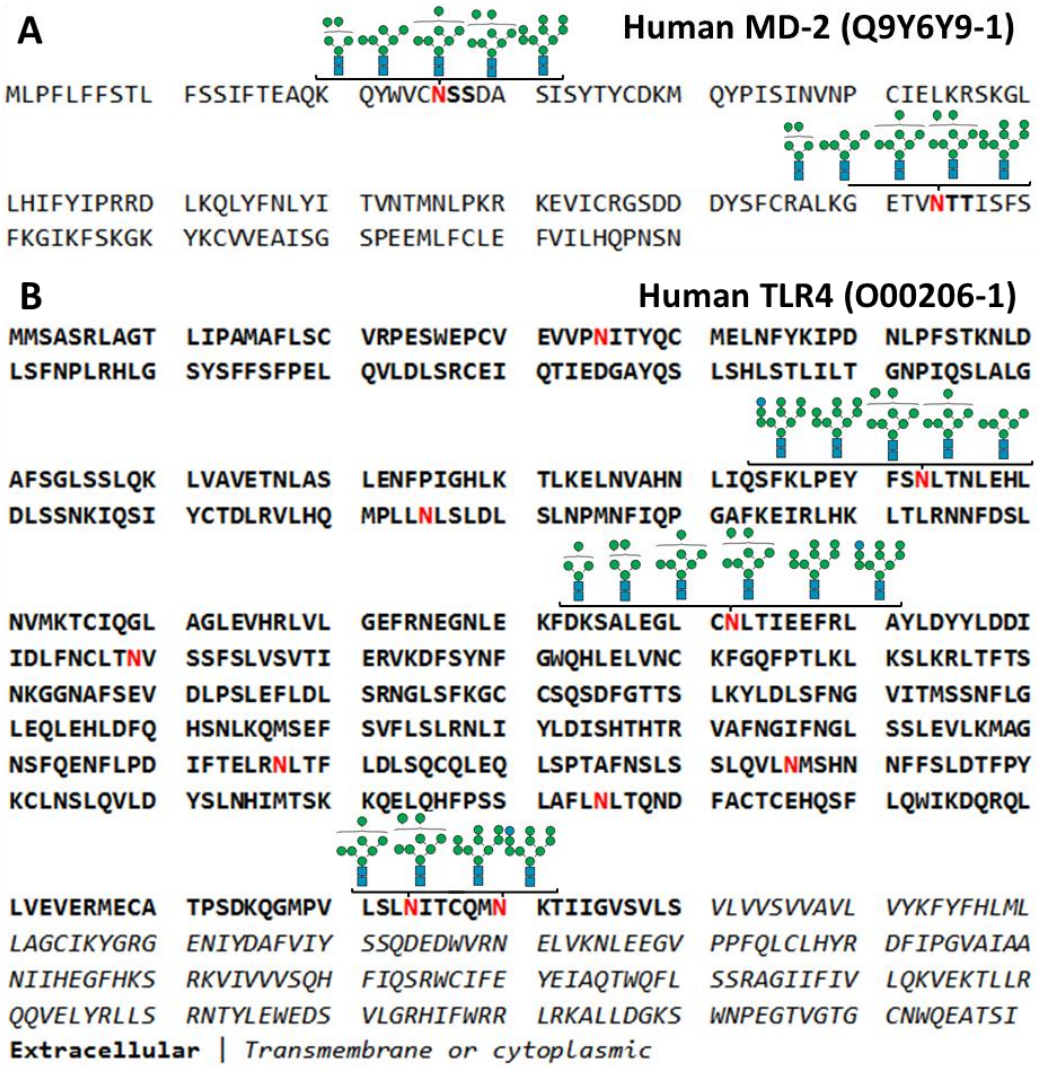


Figure 8.5 Glycopeptide coverage of HEK293-derived recombinant proteins, MD-2 and TLR4. Possible *N*-glycosylation sites, based on motif, highlighted in red.

As shown in Fig 8.6D, only high mannose *N*-glycan structures were observed for MD-2. As shown in Fig 8.5A and B, the smallest observed glycan on MD-2 was $(Man)_5(GlcNAc)_2$ with $(Man)_6(GlcNAc)_2$ as the most predominant glycan for both *N*-glycosylation sites. As two bands were cut out of the gel for recombinant human MD-2 from HEK293 cells, the *N*-glycopeptide relative abundances were then compared.

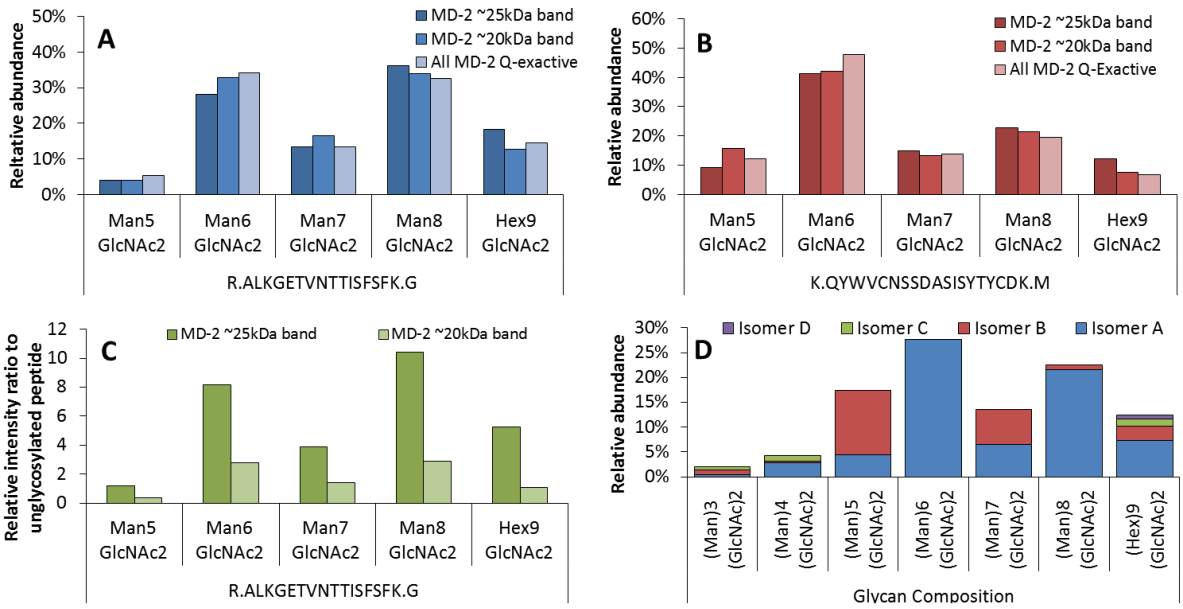


Figure 8.6 Glycopeptide and glycan analysis of recombinant HEK293-derived MD-2. Panels A and B: Glycopeptide micro-heterogeneity for given *N*-glycopeptide sequences. Panel C: Qualitative comparison of glycopeptide abundance compared to unglycosylated peptide for a given excised band. Panel D: Released glycan analysis used for populating glycopeptide search space and identifying glycan structure isomers, glycan abundances displayed are additive.

Overall, there were no significant differences in glycopeptide micro-heterogeneity between MD-2 from different gel bands. When considering the relative intensity of the detected glycopeptides compared to the unglycosylated peptide (representing macro-heterogeneity), the band of ~20kDa had a lower ratio. We suspected that the larger gel band 3 (~25kDa) corresponds to MD-2 with a greater amount of *N*-glycosylation which therefore has a larger molecular weight. Unfortunately, with the bottom-up proteomic approach that we undertook, whole protein *N*-glycosylation site occupancy could not be determined but our SDS-PAGE results suggest the presence of at least two proteoforms of recombinant MD-2 that could be separated at ~20kDa and ~25kDa.

For MD-2, the observation of only high mannose *N*-glycans was especially curious as it has been previously reported to be sialylated[230]. For glycoproteins, it has been reported that the expression and purification method for recombinant glycoproteins can result in altered glycosylation, not representative of the glycoforms observed in typical biological conditions[275]. As the “purified” recombinant protein was in a mixture with tubulin contamination (Supp Table 8.2) and only decorated with high mannose *N*-glycans, representing only the early glycosylation pathway, these

recombinant glycoproteins may not be representative of their mature forms. Despite this, we have provided the first data on *N*-glycosylation occupancy of MD-2 and TLR4 expressed by HEK293 cells using bottom-up proteomics.

8.2.3 Building a target peptide list for target TLR4 and MD-2 detection

Targeted proteomics methods typically involve the selection of specific *m/z* windows for precursor and/or product ion detection, discarding ions not in the pre-defined inclusion lists[223]. The accumulation of these *m/z* targets serves as a useful post-injection enrichment of target peptides to confirm their presence in complex mixtures. As we have been unable to detect TLR4 and MD-2 using untargeted strategies, this post-injection enrichment could improve our chances of their detection.

To identify the specific *m/z* windows to be used for targeted detection of TLR4 and MD-2, we considered several parameters for optimal peptide selection (Table 2)[219]. By taking these considerations into account, the specificity (peptide length >4 amino acids) and detectability (preferably no amino acids that can feature variable modifications) of peptides used for targeted TLR4 and MD-2 detection can be improved.

Targeted proteomics peptide selection criteria	**TLR4**	**MD-2**
No missed cleavages and peptide length >4 amino acids	32	7
AND no cysteine	22	4
AND no methionine	21	2
AND no NxS/T (*N*-glycosylation motif)	19	1
AND observed in recombinant protein	13	1

Table 8.2 Selection criteria for developing a targeted proteomics method

Applying these considerations to TLR4 and MD-2, we built a list of peptides to be used for a targeted proteomics method (Table 3). To begin the targeted method development, we used a Sciex 5600 QTOF with the same instrument parameters as typical shotgun proteomics experiments[276] except for the use of an MRM-HR method, using the inclusion list from Table 3. This method, while specific for only precursor masses, subsequently scans the full mass range for product ions for the given precursor ion[65].

Protein	Peptide	*m/z*	Molecular ion
MD-2	K.GLLHIFYIPR.R [58, 67]	410.2449	$[M+3H]^{3+}$
	K.GETVNTTISFSFK.G [109, 121]	477.5769	$[M+3H]^{3+}$
	R.GSDDDYSFCR.A [96, 105]	611.2275	$[M+2H]^{2+}$
		407.8207	$[M+3H]^{3+}$
	K.QLYFNLYITVNTMNLPK.R [72, 88]	696.6993	$[M+3H]^{3+}$
		522.7763	$[M+4H]^{4+}$
	K.MQYPISINVNPCIELK.R [39, 54]	967.9894	$[M+2H]^{2+}$
		645.6620	$[M+3H]^{3+}$
TLR4	K.NLDLSFNPLR.H [57, 66]	594.8222	$[M+2H]^{2+}$
		396.8839	$[M+3H]^{3+}$
	K.IPDNLPFSTK.N [47, 56]	566.3059	$[M+2H]^{2+}$
	K.TCIQGLAGLEVHR.L [244, 256]	485.2575	$[M+3H]^{3+}$
	K.GCCSQSDFGTTSLK.Y [388, 401]	774.3269	$[M+2H]^{2+}$
	K.ELNVAHNLIQSFK.L [153, 165]	504.9438	$[M+3H]^{3+}$
	K.IQSIYCTDLR.V [186, 195]	634.8188	$[M+2H]^{2+}$
	K.LVAVETNLASLENFPIGHLK.T [130, 149]	722.4037	$[M+3H]^{3+}$
		542.0546	$[M+4H]^{4+}$
	R.QQVELYR.L [780, 786]	468.2509	$[M+2H]^{2+}$

Table 8.3 Peptides and corresponding *m/z* values used as inclusion list for using targeted proteomic method with MRM-HR for detection of TLR4 and MD-2. Modified amino acids are underlined.

To validate the PRM method, recombinant human TLR4 and MD-2 from HEK293 cells were separated by SDS-PAGE, bands of sizes consistent with the target proteins (~110kDa and ~20kDa) digested with trypsin and C18 cleaned peptides were analysed by mass spectrometry. As shown in Fig 8.7A, MD-2 had a single, intense, predominant peptide, with very good b/y fragment ion coverage. Although this peptide did not satisfy the peptide selection criteria of no cysteine (which can introduce peptide quantitation differences depending on alkylation efficiency[277]), it was the most intense, therefore the best observed peptide for detecting MD-2 in biological samples.

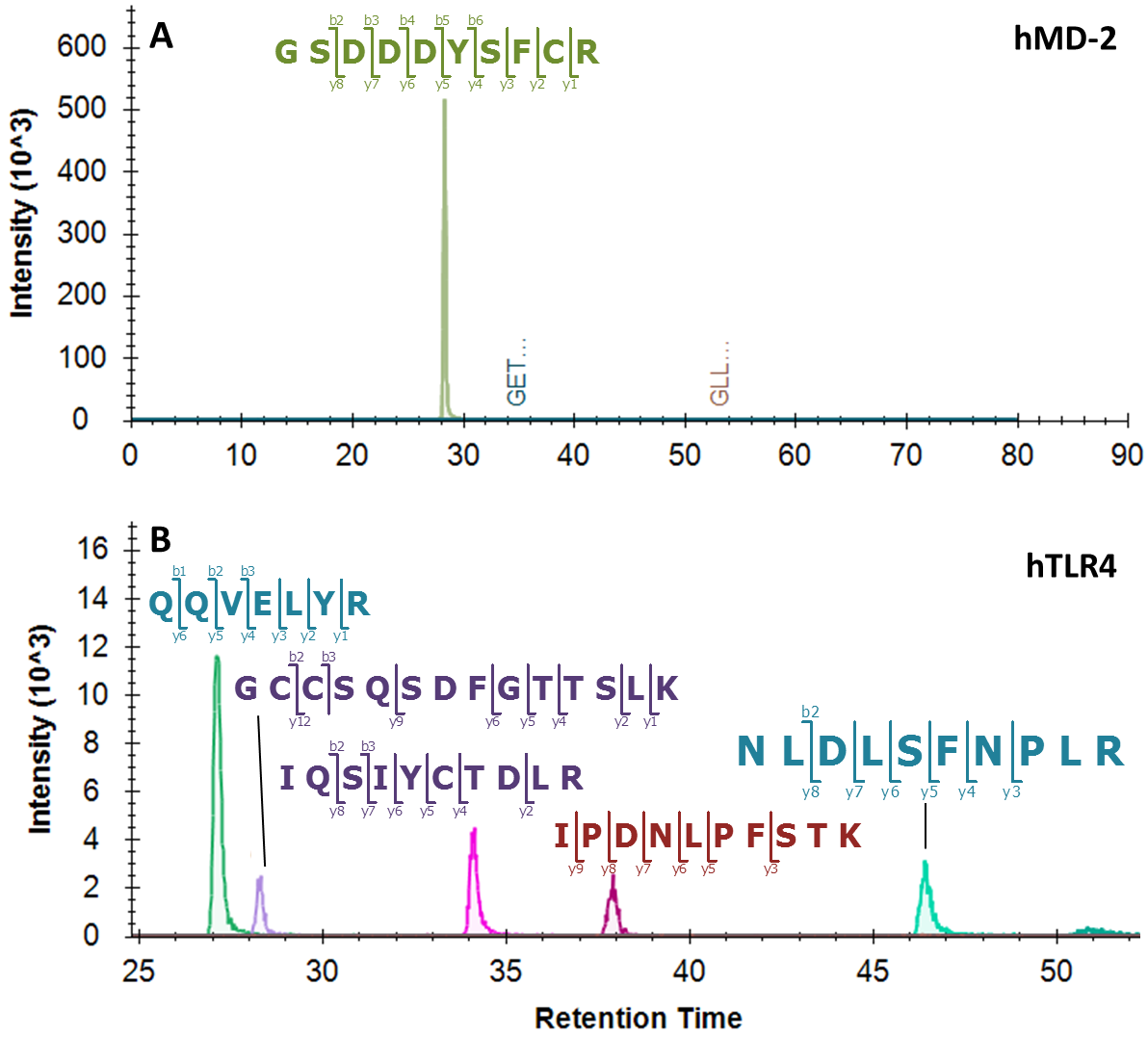


Figure 8.7 Selection of peptides for MRM-HR (aka PRM) targeted detection method with recombinant protein: top, human MD-2, and bottom, human TLR4. Peptide fragment backbone coverage denoted above corresponding peptide peak.

These results for MD-2 were contrasted by human TLR4, that had 5 intense peptides detected with good peptide backbone coverage by MS/MS and 3/5 of these peptides fulfilled all the peptide selection criteria shown in Table 2. The difference in number of good peptide targets for MD-2 and TLR4 can largely be attributed to the protein length, with the amino acid mass of MD-2 being 1/5th of the amino acid mass of TLR4; therefore it can be expected that there are approximately 1/5th of the peptide targets detected. To improve the number of peptides detected for both proteins, other proteolytic enzymes could be used such as Asp-N (cleaves C-terminal to N) or chymotrypsin (cleaves C-terminal to W, Y, F) which are frequently used when trypsin does not achieve suitable coverage[278]. however we continued to use trypsin because it generates proteotypic peptides easily detected by positive mode LC-MS because of the high likelihood of containing positively charged C-terminal lysine and arginine residues[279].

Following selection of suitable peptides for the PRM method using recombinant human TLR4 and MD-2, protein extracts of subcellular fractions from one million HEK293 cells with and without stable transfection of TLR4, MD-2 and CD14, were separated by SDS-PAGE (Fig 8.8). In addition to the previously used chemical subcellular fractionation, digitonin-based method (as used in Fig 8.1B), an ultracentrifuge based method was used to obtain a total membrane subcellular fraction, in the hope of another opportunity for enrichment of TLR4. While the chemical subcellular fractionation method theoretically achieves membrane enrichment through chemical cell perforation and subsequent membrane solubilisation, the ultracentrifuge-based physical membrane enrichment method exploits the different density between cytoplasmic soluble and plasma membrane proteins. Some protein extracts were run in duplicate, to compare molecular size-specific and total lane gel excision and digestion approaches.

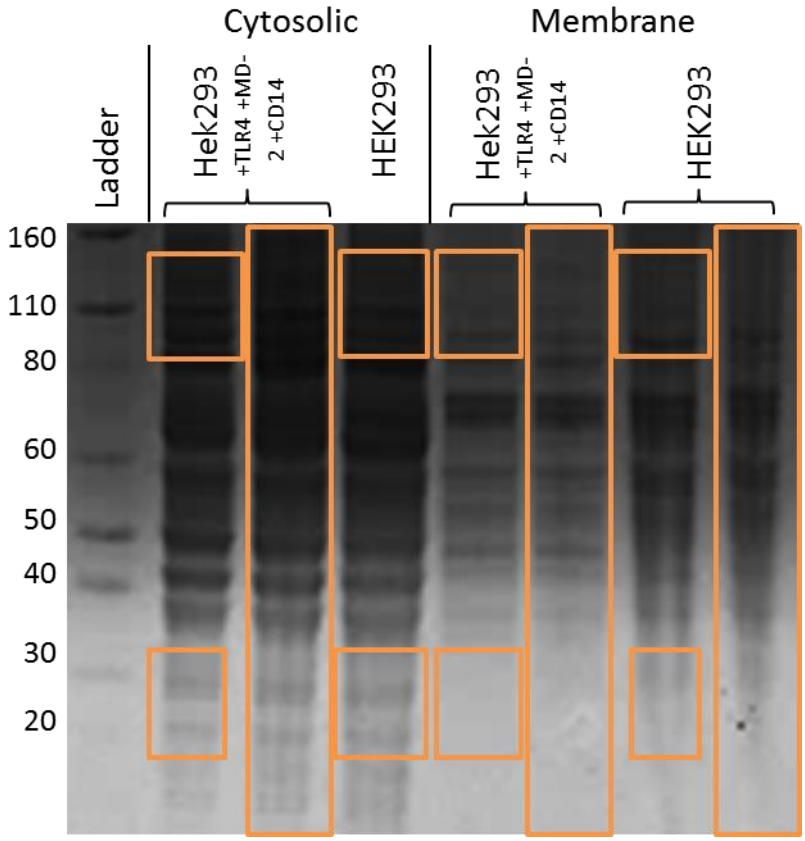


Figure 8.8 SDS-PAGE of extracted proteins from digitonin-based subcellular fractionation of HEK293 and stably transfected cell lines. Highlighted boxes represent excised regions for subsequent targeted proteomic analysis.

The gel regions were trypsin digested after excision, C18-cleaned and analysed with the selected peptides in the targeted proteomics method using C18-LC-ESI-MS/MS. Surprisingly, no TLR4 or MD-2 peptides were detected in any of the samples.

8.2.4 Using an alternate mass spectrometer for more sensitive targeted proteomics

As TLR4 or MD-2 were not detected in membrane enriched samples of a stably transfected cell line confirmed to translate the representative genes, we considered possible areas for increasing the sensitivity of detection. As we were utilising instruments as part of a core-facility, using more than 1µg of material was not possible due to the limited capacity of the trap and LC-MS set up. One improvement available to us was to use a quadrupole-Orbitrap based instrument (Thermo Q-Exactive), to allow for increased product ion intensity resulting from precursor accumulation which cannot be performed on a QTOF-based instrument (Sciex 5600). We analysed the same recombinant proteins (TLR4 and MD-2) on the quadrupole Orbitrap using the shotgun proteomics method to modify the inclusion list for the targeted proteomics method used on with the QTOF. As shown in Figure 8.9, peptide intensity and peptide backbone fragmentation coverage were approximately the same, if not better, than those observed on the QTOF. For recombinant human TLR4, the relative abundance of the peptides changed, and LTFTSNK which was not seen by the QTOF, was detected as the second most abundant peptide. No other peptides were observed for recombinant human MD-2.

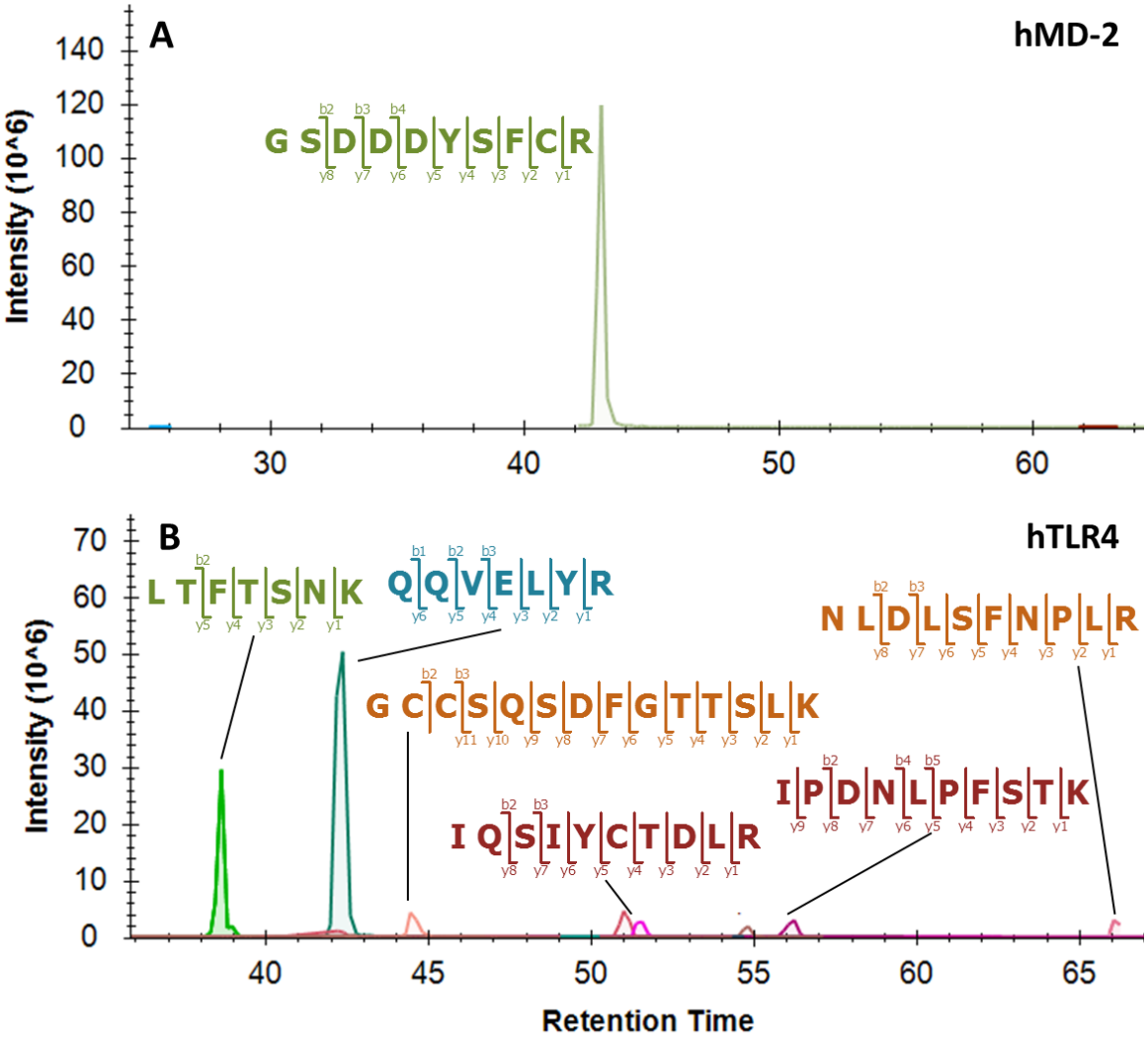


Figure 8.9 Analysis of recombinant proteins with an Orbitrap mass-analyser: top, human MD-2, and bottom, human TLR4. Peptide fragment backbone coverage denoted above corresponding peptide peak.

Based on these results, a PRM method was created with a limited inclusion list (Table 4). This inclusion list was used to allow high precursor accumulation (from 60ms to 100ms maximum accumulation time) to give a better chance of observing essential peptide backbone fragment ions for confident detection of the targeted TLR4 and MD-2 peptides.

Protein	Peptide	*m/z*	Molecular Ion
MD-2	K.GLLHIFYIPR.R [58, 67]	410.2449	$[M+3H]^{3+}$
	R.GSDDDYSFCR.A [96, 105]	611.2275	$[M+2H]^{2+}$
TLR4	K.IPDNLPFSTK.N [47, 56]	566.3059	$[M+2H]^{2+}$
	R.QQVELYR.L [780, 786]	468.2509	$[M+2H]^{2+}$
	R.LTFTSNK.G [355, 361]	405.7214	$[M+2H]^{2+}$
	R.QLLVEVER.M [598, 605]	493.2875	$[M+2H]^{2+}$

Table 8.4 Peptides and corresponding *m/z* values used as inclusion list for targeted proteomic method with PRM on the Q-Exactive

In addition to changing instruments for analysis, we stimulated the HEK293 cell lines with LPS, an agonist of the TLR4 pathway. As shown in Fig 8.10, cells with and without LPS treatment were subjected to chemical subcellular fractionation and protein extracts analysed by SDS-PAGE (methods described in subchapter 2.4). Gel regions at the expected masses of TLR4 and MD-2 were excised and peptides prepared for LC-MS/MS as previously.

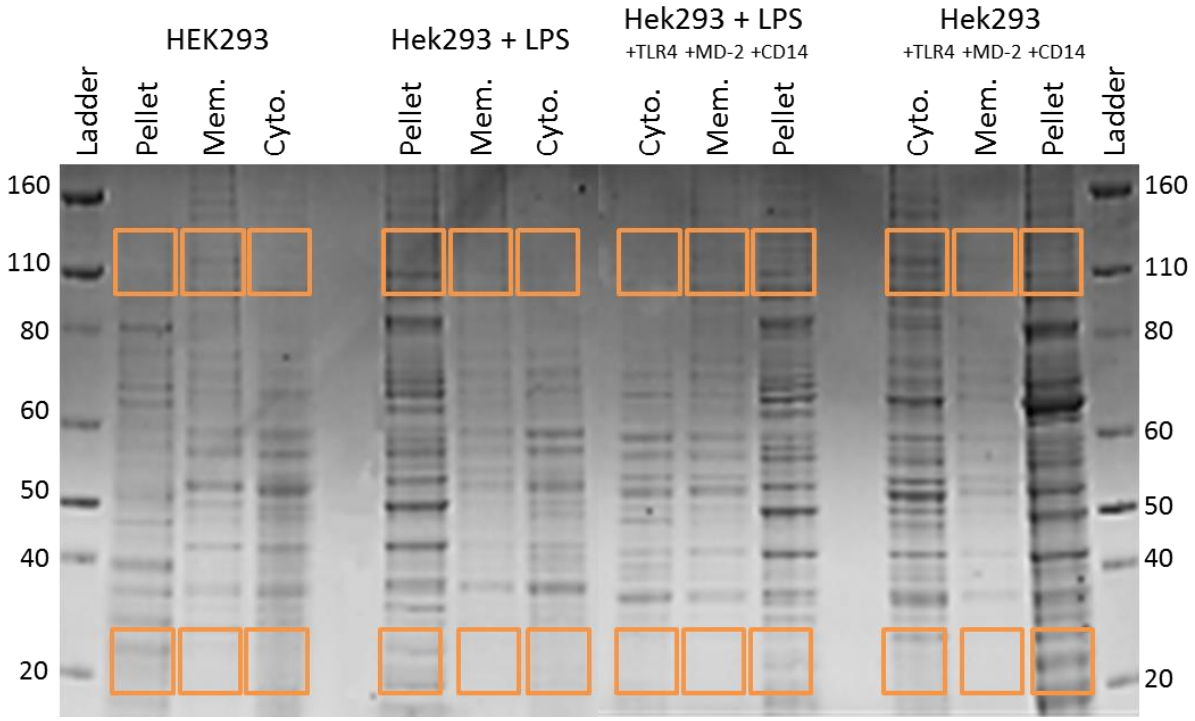


Figure 8.10 SDS-PAGE of extracted proteins from subcellular fractions of HEK293 and stably transfected cell lines treated with or without LPS. Highlighted boxes represent excised regions for subsequent targeted proteomic analysis.

Following sample analysis with our PRM method, a single peptide from human MD-2 was detected (GSDDDYSFCR) in the cytosolic protein fraction of the transfected HEK293 cell line, which was also the best ionising peptide from analysis of recombinant MD-2. Figure 8.11 shows the abundances of the peptide backbone y ions with this peptide detected in the cytosolic subcellular fraction of HEK293 cell lines stably transfected with TLR4/MD2/CD14 with the combined product ion abundance of 0.1% of peptide backbone ions detected from 1µg of recombinant MD-2. We initially had concerns regarding sample carryover being responsible for this detection however the samples immediately before each biological sample did not detect GSDDDYSFCR, affirming the detected peptide originated from the cytosolic protein fraction of the transfected HEK293 cell line. Although it was not possible to detect MD-2 in its mature secreted form in the culture media due to the contamination due to using 10% (v/v) bovine serum in the growth medium, the detection of MD-2 in a cytosolic subcellular fraction is consistent with its translocation from the cytoplasm to be secreted from the cell[280]. The detection of MD-2 is interesting as MD-2 assists in the translocation of TLR4 to the cell surface and

without expression of MD-2 protein, TLR4 resides in the Golgi apparatus[280, 281]. However, no TLR4 targeted peptides were detected in any cellular fraction.

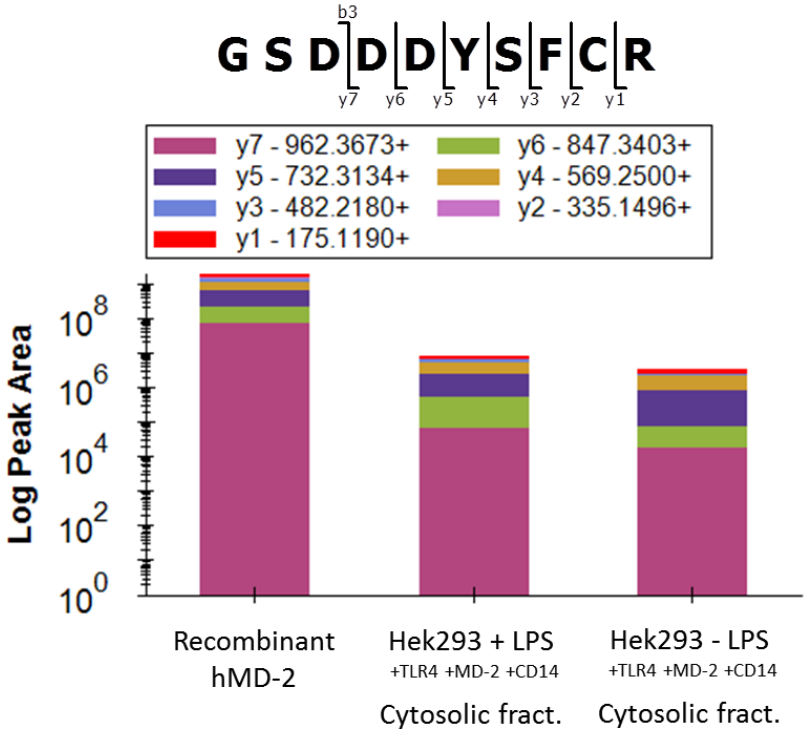


Figure 8.11 Targeted PRM proteomic analysis of MD-2 excised gel bands from Fig 8.4.

8.2.5 Understanding why TLR4 could not be detected

To determine the limit of detection of TLR4 by PRM and to investigate why TLR4 was not detected in the transfected cells, we generated a standard dilution curve for TLR4 and MD-2 with and without cell lysate to emulate the conditions in which previous analyses were conducted, where we were looking for TLR4 peptides in a cell lysate or membrane protein fraction (Fig 8.12). From 125fmol to 1250fmol, all targeted TLR4 peptides gave a linear response with a wide range of intensities on the mass spectrometer, varying from 2.5e9 to 9.2e7 monoisotopic precursor peak area for five TLR4 peptides, which is a 27-fold difference. This highlights the discrepancy between the responses of equimolar peptides analysed by mass spectrometry due to properties that influence ionisation efficiency, such as hydrophobicity and charge[282].

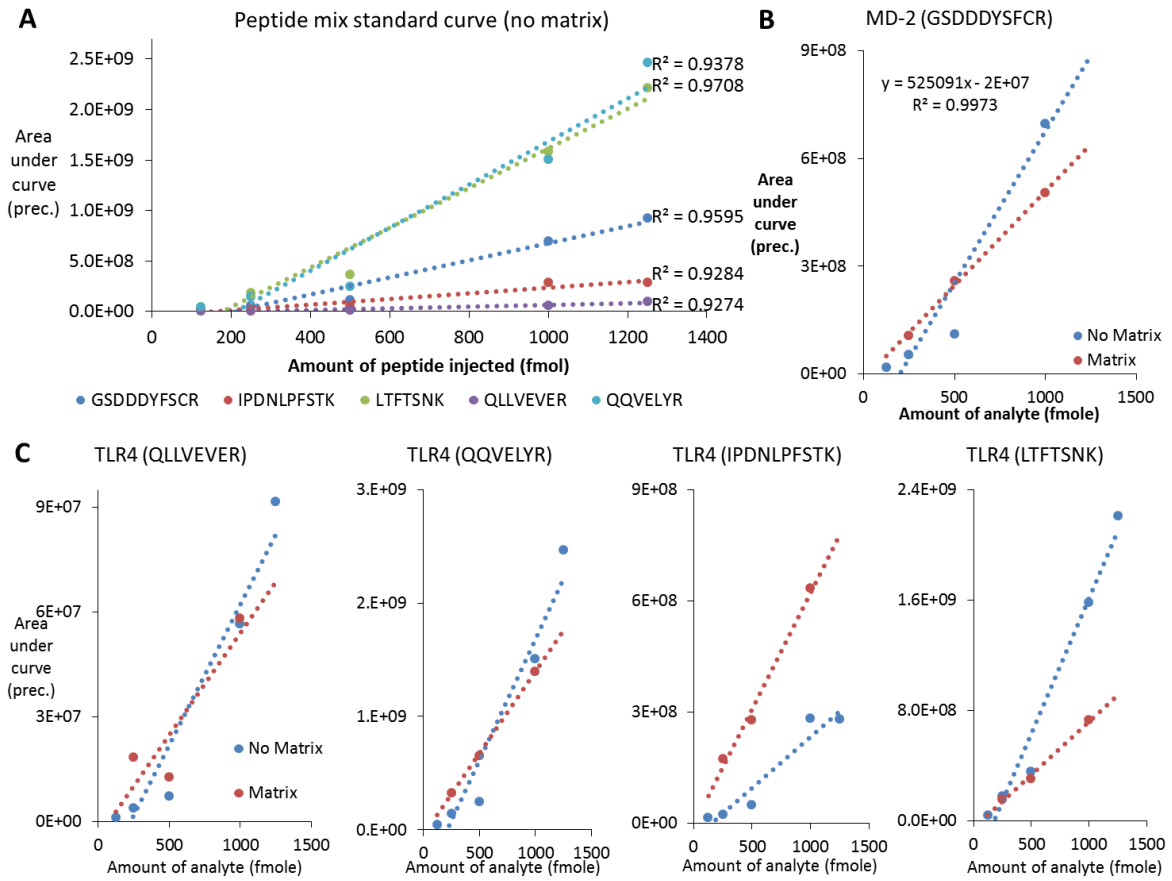


Figure 8.12 Dilution ionisation curves for peptides used in the targeted method for TLR4 and MD-2 detection with/without added complex peptide matrix. A) Dilution curve for each peptide based on peptide precursor area, (B) Standard curve for most intense peptide of MD-2. (C) Dilution curve for the most abundant peptides of TLR4. Standard curves shown in red contain added peptide matrix from a HEK293 cell lysate. Precursor is shortened to Prec.

Below 500 fmol for all peptide targets, validation of peptide sequence was difficult (>3 y ions), resulting in one peptide (QLLVEVER) having a limit of detection at 250fmol and the remaining peptides for TLR4 and MD-2 having limits of detection at 125fmol when analysed with no cell lysate matrix. A summary of the limit of detections and linearity ranges can be found in Table 4. Both TLR4 and MD-2 peptides, GSDDDYSFCR and IPDNLPFSTK respectively, had the lowest limit of detection at 25fmol. The peptide samples containing the cell lysate matrix samples were prepared to have double the concentration of peptides for each dilution series due to expected suppression of our peptides of interest by peptides from the cell lysate sample. Surprisingly, the cell lysate matrix had little effect or in some cases, increased the precursor area corresponding to peptides of interest, such as the for TLR4 peptide IPDNLPFSTK. This may also be attributed to interfering matrix when quantifying based on precursor instead of product ions. To reduce the influence of the matrix, product ion-based quantitation was performed but was found to have less linearity (Supp Fig 8.1).

Protein	Peptide sequence	No matrix			Cell lysate matrix		
		Limit of detection (fmol)	Linearity range	R^2	Limit of detection (fmol)	Linearity range	R_2
MD-2	GSDDDYSFCR	125	125-1250	0.96	25	25-2000	1.00
TLR4	IPDNLPFSTK	125	125-1250	0.93	25	25-2000	0.94
	LTFTSNK	125		0.97	250	250-2000	0.99
	QLLVEVER	250		0.94	250		0.91
	QQVELYR	125		0.94	25	25-2000	0.99

Table 8.5 Limit of detection statistics for the most abundant peptides of TLR4 and MD-2 with no matrix and with peptide matrix from HEK293 cell lysate.

8.2.6 Looking to alternate sample sources for TLR4 and MD-2 detection

As TLR4 could not be detected from the membrane enriched fraction of one million stably transfected HEK293 cells (with a method limit of detection of 25 fmol), we looked for other cell lines that may express TLR4. The U87MG human glioblastoma cell line was chosen as it is frequently used for investigation into the TLR4 signalling pathway and has been confirmed to express mRNA for TLR4, MyD88 and CD14[283]. Also, in an attempt to detect secreted MD-2, we grew the U87MG cell lines, as well as the HEK293 cells stably transfected with TLR4/MD-2/CD14, treated with either an agonist (LPS) or inhibitor (LPS-RS) in serum-free media and harvested the media after the cells were 70% confluent. The serum-free media was taken at pre-defined time points (0 to 44 hours) and analysed by SDS-PAGE (Fig 8.13A). Following removal of the serum-free media, subcellular fractionation was performed and extracted proteins were analysed with SDS-PAGE (Fig 8.13B). Gel regions, where TLR4 or MD-2 were expected, were excised and prepared for targeted proteomic analysis.

We hoped that the use of a more biologically relevant cell line, human glioblastoma, would be a more suitable model compared to stably transfected HEK293 cells. However, we could not detect

human TLR4 or MD-2 in the human glioblastoma cell line whereas MD-2 alone was detected in the HEK293 stably transfected with MD-2. Aside from differences between cell lines, the use of serum-free medium could also have affected protein expression on the cell surface, as was the case for mesenchymal stem cells which only differentiated into smooth muscle cells in the presence of bovine serum[284].

Our inability to detect TLR4 is not completely surprising given its important role as one of the gatekeepers in the human immune response[285]. Over-expression of TLR4 could lead to an increased immune response, which could be potentially fatal in-cases of over-stimulation[152] but, as TLR4mRNA increases are often reported [283] following immune stimulation, TLR4 protein expression appears not to be directly related.

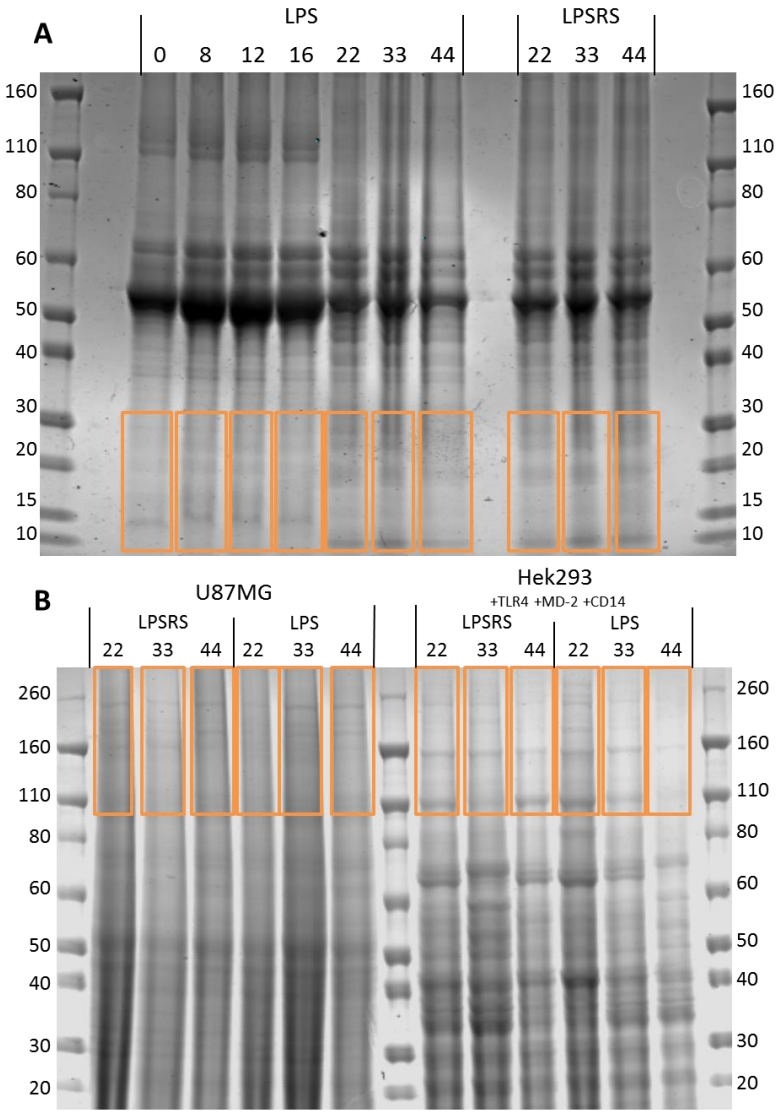


Figure 8.13 SDS-PAGE of extracted proteins from U87MG and stably transfected HEK293 cells treated with LPS or LPS-RS for defined times in serum-free media. Highlighted boxes represent excised regions for subsequent targeted proteomic analysis. Top, MD-2 region of extracted proteins from serum-deprived media from U87MG cells. Bottom, extracted proteins from membrane subcellular fractions from U87MG and stably transfected HEK293 cells.

8.2.7 Mining published proteomics datasets to determine frequency of detection of TLR4 and MD-2

To understand our difficulty in detecting TLR4/MD-2 in the stably transfected cell line and microglia cells, we mined the MS/MS data in PRIDE, a proteomics data repository, to see how frequently TLR4 and MD-2 have been detected in human samples studied with shotgun proteomics[226]. This information could then be used to inform our approach for detecting TLR4/MD-2.

One thousand human shotgun proteomic studies were found in the PRIDE protein MS/MS database, encompassing a wide analysis range of sample types and instrument types, with only 43 and 26 reporting TLR4 and MD-2 detection, respectively. Further filtering of these lists for peptides with sequences specific for TLR4/MD-2 (BlastP search against human NCBI non-redundant database[286]), high quality spectral matches (as reported by PRIDE) and tryptic peptides, this list quickly reduced to only 3 and 8 datasets reporting TLR4/MD-2 detection based on these criteria, respectively (Table 6).

Criteria	TLR4	MD-2
Human datasets	1000	1000
AND reported protein target detection	43	26
AND peptide sequence specific for target protein	32	26
AND removal of low quality spectral matches (as assigned by PRIDE)	15	12
AND 2 or more peptides (TLR4 only)	6	N/A
AND only tryptic peptides (<2 missed cleavages)	3	8
Datasets confidently detecting target protein	0.3%	0.8%

Table 8.6 Analysis of datasets on PRIDE for confident human TLR4 and MD-2 detection

To understand these detected TLR4 peptides, coverage of the full human TLR4 sequence from the three PRIDE datasets (PXD002462, PXD001974 and PXD003924) was mapped onto the protein sequence representing the most common isoform, isoform1A (Fig 8.14). Two peptide sequences were detected in all three datasets (QGMPVLSLNITCQMNK and QQVELYR), which together, represent peptide detection at the intracellular and extracellular domains of the protein. Despite these proteins not being mentioned in the publications accompanying these datasets[287–289], the accompanying experimental methods for these studies provides a valuable insight into the requirements for TLR4 and MD-2 detection with shotgun proteomics as described below.

Figure 8.14 Map of PRIDE datasets reporting confident human TLR4 detection. PXD###### represents the PRIDE accession number. Peptide sequences shared across all reported datasets highlighted in red.

All of the studies that were confident of the detection of human TLR4 were using cell lines (HeLa, HEK293 and RCMH). Interestingly, one study features a HEK293 cell line, similar to the cell line control we used for our shotgun proteomic methods with the detection of a single TLR4 peptide sequence. This conflicts with our experimental evidence demonstrating that HEK293 cell lines do not express the necessary human TLR4 mRNA, which is supported by an earlier publication with the same finding[273]. To prevent issues such as these, a cell-line-specific FASTA databases based on RNA-seq data should be used to ensure protein identifications are consistent with mRNA expression in the cell line of interest[290].

In addition to the protein identification lists generated by these range of studies in PRIDE, each project also documents the experimental methods used to generate the protein identification lists, thus allowing us to identify the parameters used to obtain the three datasets that reported TLR4 detection. A summary of the similarities and differences in the experimental methods are compared in Table 7. Overall, all workflows have fractionation prior to MS analysis with size, affinity or C18 fractionation, as well as using different amounts of starting material, with one study starting with 2mg of cells. In addition, all studies utilise mass-spectrometers capable of trapping ions (linear ion trap or Orbitrap) with a lowest total acquisition time of 35 hours. These resource requirements for TLR4 detection with shotgun proteomics make the prospect of repeated TLR4 detection for quantitative applications challenging.

PXD # (or method used)	Protein Source	Sample preparation (starting protein weight)	MS used	Number of raw files in PRIDE project and length of gradient (min. acquisition time)	Ref.
PXD002462	Human embryonic cell lysate (HEK293T)	Affinity purification, (not mentioned)	LTQ Velos ion-trap (Thermo)	88 runs x 30 min (44 hrs)	[287]
PXD001974	Human myoblast cell lysate (RCMH)	C18 fractionation @ pH 6.0 (~2mg)	LTQ Orbitrap elite (Thermo)	16 runs x 130 min (35 hrs)	[289]
PXD003924	Secreted proteins (HeLa)	SDS-PAGE size fractionation (~540ug)	LTQ Orbitrap (Thermo)	54 runs x unspecified	[291]
Our shotgun method	Cell lysates (Hek293, U87MG)	SDS-PAGE size fractionation (~20ug)	LTQ Q-Exactive (Thermo)	1 run x 90 min	
Our targeted method	Cell lysates (Hek293, U87MG)	SDS-PAGE size fractionation (~20ug)	LTQ Q-Exactive (Thermo)	1 run x 60 min	

Table 8.7 Metadata of the PRIDE datasets reporting TLR4 peptide detection. For peptide sequences corresponding to each PXD number, please see Figure 8.14.

We then evaluated the raw data for the most successful dataset that detected human TLR4; PXD001974, which reported detection of 6 tryptic TLR4 peptides. Using search files (.mzid) for the TLR4 peptides from the raw files, we extracted the exact MS/MS spectra that were reported to match the reported TLR4 peptides (*m/z* value, sequence and any modifications) for manual inspection. As shown in Figure 8.15, following inspection of the data with Skyline protein identification software, most of the peptides detected for TLR4 were not convincing peptide identifications. For example, for the topmost peptide, NLDLSFNPLR, while predicted abundant b and y ions were in the MS/MS spectrum, the missing monoisotopic peak means the peptide mass matching the unmodified TLR4 amino acid sequence, NLDLSFNPLR, was not actually detected.

Digging deeper into the search parameters used in this study showed that the authors allowed off-by-one errors in precursor *m/z* value for peptide sequencing matches. These off-by-one errors are typically encountered with large peptides where the more-abundant second isotopic peak is selected for MS2[292]. In this case, the off-by-one error was instead caused by deamidation (+0.98 Da) of the first asparagine in the peptide sequence. This error can remain unnoticed as abundant y-ions were detected based on the unmodified peptide sequence, as they are not impacted by the deamidated asparagine. In addition, the isotope distribution of this peptide has the monoisotopic peak as the most abundant. When the correct peptide sequence was searched, N[+0.98]LDLSFNPLR, the isotope distribution and MS2 spectra confirmed its detection.

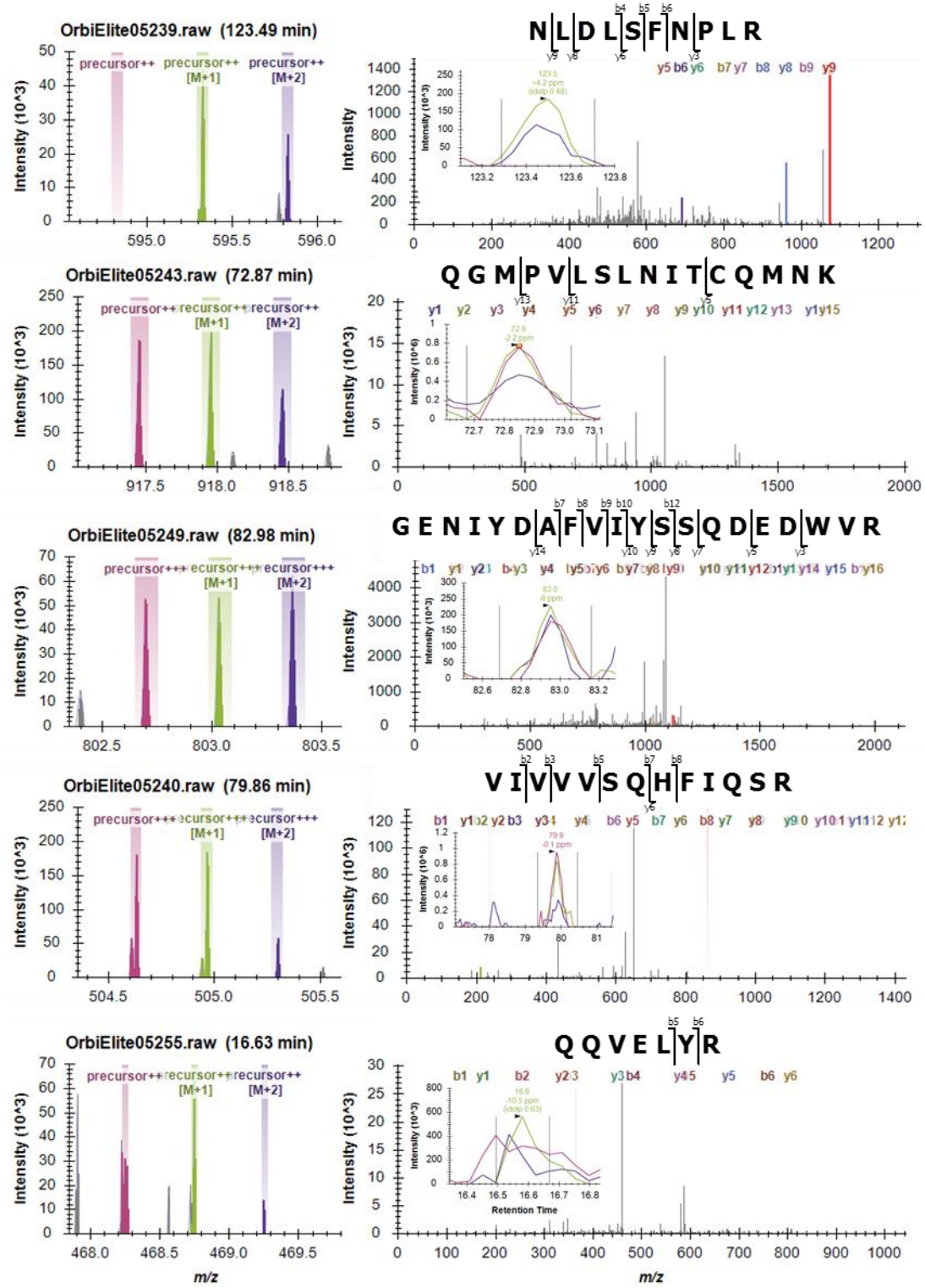


Figure 8.15 Raw data re-analysis of published PRIDE dataset PXD001974 which detected the greatest number of TLR4 peptides. On the left, MS1 spectrum closest matching peptide of interest with expected isotopic distribution highlighted. Right, closest matching MS/MS mass spectrum for the given peptide with detected fragments annotated and spectrum highlighted in colour. Centre inset, isotopic overlap, retention time and mass accuracy error for given peptide.

For the remaining peptides in Fig 8.15, all the expected isotopes were present, at the correct charge state and with sufficient mass accuracy (<10 ppm) however, upon inspection of the MS2 spectra cited as TLR4 peptide sequence matches, the low relative abundance of the fragment ions casts

doubt on the reported TLR4 detection in this dataset. This initially promising dataset reduced our confidence in the reported detection of TLR4 in the two remaining PRIDE datasets.

In the same way, mapping 8 proteomic datasets that reported confident MD-2 detection in PRIDE; only 1 peptide for MD-2 was detected in each (Fig 8.16). To confirm MD-2 detection at least 2 peptides should be detected. The peptide sequence QLYFNLYITVNTMNLPKR was detected by four of the datasets but surprisingly, the most intense peptide observed with our targeted proteomic analysis of trypsin-digested recombinant MD-2 (GSDDDYSFCR) was not observed.

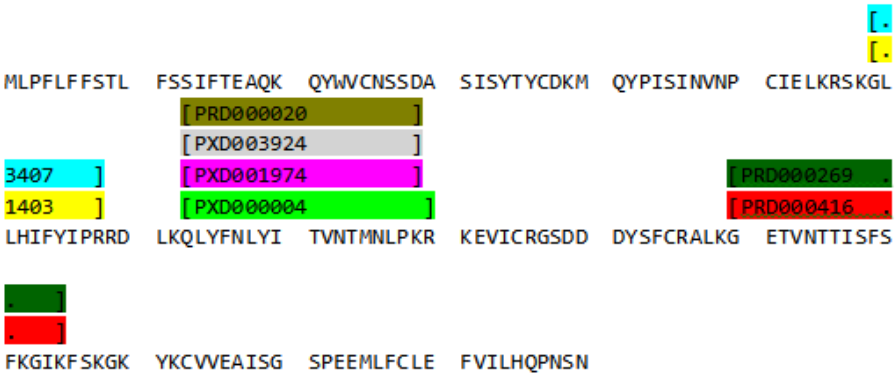


Figure 8.16 Map of PRIDE datasets reporting confident human MD-2 detection

Considering PRIDE database results for both proteins, only 2 datasets reported TLR4 and MD-2 detection. One of these datasets, PXD001974, was found to be somewhat unreliable in its reporting of TLR4 detection after manual inspection of MS1 and MS2 spectral information for each reported peptide (Fig 8.15). Our search found that TLR4 and MD-2 are not frequently detected in shotgun proteomics experiments, and if they are found, they require extensive fractionation of samples from plentiful starting material.

8.2.8 Where in the world is TLR4, and how much is there?

The inability to detect human TLR4 protein in cell lines, even with a targeted proteomics method, is concerning given its expected integral role in the human immune system[293]. In the previous chapter, we confirmed transcription of TLR4 mRNA in our stably TLR4 transfected HEK293 model cell line and published research has confirmed TLR4 transcription in this same cell line and in the U87MG human glioblastoma cell line[283]. We utilised LPS as an agonist to hopefully increase TLR4 protein expression as TLR4 mRNA is upregulated approximately 3-fold between 2-8 hrs following TLR4 pathway activation by LPS[294]. Despite using different cell lines, agonists and targeted mass-spectrometry, we could not detect TLR4 in any samples and our analysis of stored proteomics datasets reporting TLR4 detection revealed that these identifications are few and are likely not of high confidence.

Based on our LOD determination for TLR4, 25fmol of TLR4 peptides are required for detection which represents approximately 1500 copies of TLR4 per cell (from our starting material of 1 million cells). A dataset performing absolute protein quantification in the U87MG cell line identified 237 out of 6352 total detected proteins with fewer than 1500 copies, putting it in the bottom 4% of all proteins detected by copy number[295].

8.3 SUMMARY AND CONCLUSIONS

Despite using a HEK293 cell line stably transfected with CD14. TLR4 and MD-2, and confirming transcription of their genes, these proteins could not be detected using shotgun proteomics. A shotgun proteomic study of the TLR4 pathway under LPS stimulation in these cells also could not detect the secreted and membrane-bound front line of the pathway, demonstrating that a shotgun proteomic study was not suitable.

Analysis of recombinant human MD-2 and TLR4 identified high mannose glycan structures to be on these proteins that were expressed in HEK293 cells. With moderate protein sequence coverage, this provided a starting point for the development of a strategy to detect these glycoproteins in biologically relevant samples. A targeted method for TLR4 and MD-2 detection was established on QTOF and Q-Orbitrap platforms with *E. coli* and these HEK293 expressed recombinant proteins were used to optimise the targets and validate the methods. Proteins from subcellular fractions of a HEK293 cell line stably transfected with TLR4 and MD-2 were analysed with these methods. A single peptide of MD-2 was detected in the cytosolic fraction of these cells and no TLR4 peptides were detected.

In attempts to increase the abundance of TLR4 and MD-2, the transfected HEK293 cells were treated with LPS but again, only a single peptide from MD-2 and no peptides from TLR4 were detected. Standard curves of diluted peptides from TLR4 and MD-2 were made, identifying the limit of detection for the peptide targets that allowed detection down to 25fmol for both TLR4 and MD-2 proteins, representing approximately 1500 protein copies per cell.

Analysis of previously published human proteomic datasets in the PRIDE MS/MS proteomics database, showed confident detection of TLR4 and MD-2 was achieved in only 0.3% (3/1000) and 0.8% (8/1000) of these datasets. Manual analysis of the spectra from the dataset with highest reported coverage of TLR4, revealed that these reported peptides did not feature sufficient evidence for confident peptide assignment (fragmentation spectra and correct precursor mass), except for a single observed peptide which was reported incorrectly.

TLR4 protein detection has been reported with antibodies, leaving two significant possibilities: the antibodies are more sensitive than targeted mass-spectrometry or that the reported TLR4 detections are by off-target antibodies. While antibodies can certainly be more sensitive than mass-spectrometry[296], poor antibody specificity is also a well-known risk for antibody-based assays[245, 246]. In regards to where TLR4 is, and how much is there, surprisingly SDS-PAGE separation of proteins followed by targeted mass-spectrometry appears to be unable to answer these questions.

CHAPTER 9 – SUMMARY AND FUTURE DIRECTIONS

9.1 SUMMARY

In this study, we have developed and improved knowledge in two areas: method development for LC-MS-based glycomics and the TLR4 pathway. In chapter 3, we demonstrated that optimisation of vendor-supplied MS parameters and methods are essential for quantitative performance in the field of glycomics. Although these results are system-specific, reducing the applicability of our parameters for other researchers, we provide a set of parameters worth considering for quantitative MS applications.

In the next chapter, we evaluated fragmentation parameters and found that optimisation of these is important for maximising the glycan characterisation afforded by LC-MS-based analysis. Comparing fragmentation methods, we observed that a variety of CID methods produce qualitatively different fragmentation spectra for the same glycan structure, highlighting one possible source of variation between MS/MS spectra of the same glycan structure using different variations of CID. We also found that the variation in MS/MS spectra for a single glycan structure with one technique was greater than the difference in MS/MS spectra between glycan structural isomers, identifying a possible concern with spectral library-based discrimination of glycan isomers.

This chapter also includes my first experimental research paper, identifying an acidic monosaccharide component of *O*-glycans from secreted proteins from fungi, the first time this has been reported. This identification was the result of using an alternative CID method capable of overcoming one of the fundamental limitations in ion-trap resonant CID and resulted in the characterisation of their composition, sequence and structures. Combined with other analytical techniques, monosaccharide analysis and glycoproteomics, gaps in our characterisation were covered, identifying the monosaccharide component as glucuronic acid and the *O*-glycan modification to the catalytic domain of cellobiohydrolase I.

In chapter 5, we developed a workflow for automated glycan isomer discrimination using diagnostic ions. To achieve this, we evaluated the contribution of collision energy to the production of diagnostic ions in MS/MS spectra of several glycan structures. Following this, we applied our vendor-neutral and open-source workflow for automated glycan isomer discrimination, achieving discrimination of sialic acid linkage and arm composition isomers.

The next chapter improved on the use of retention time for glycan characterisation through the development of a normalisation method using a dextran ladder as an internal standard. The dextran ladder allowed normalisation of retention time and mass spectrometer response, reducing technical variation. This normalisation of retention time was then used to build a library of glycan structures with system-independent PGC-based retention times with associated MS/MS spectra. Finally, we

developed a spectral library approach for automating peak assignment of the dextran ladder in complex, and isomeric, glycan mixtures.

For the next subsection, we focused on investigating the TLR4 pathway using mass-spectrometry with the first chapter investigating the effect of TLR4 activation on protein glycosylation. Analysing the secreted protein glycosylation profile, we observed a 30-fold increase in core-fucosylation between 16 and 33 hours after TLR4 activation in a human microglia cell line. This implicates FUT8, the enzyme responsible for the addition of core-fucose to the *N*-glycan core, for the first time in the cell response to TLR4 activation.

Moving to a mouse animal model to better represent the complexity of TLR4 activation and neuropathic pain, *N*-glycans from the brain and spinal cord were analysed, identifying abundant high mannose structures consistent with other studies of the mouse brain. Using a neuropathic pain model, we quantified *N*-glycans in these tissue sections in the hope of uncovering changes to protein glycosylation in response to TLR4 activation. Unfortunately, biological variation was so great that any meaningful changes in abundance were not significant but through the correlation of glycan abundance in the PAG to mouse behaviour, the paucimannose *N*-glycan class was identified as being positively correlated to increased sensitivity to usually painless stimuli.

To uncover more about the TLR4 pathway, we sought to detect and characterise the TLR4-complex using mass-spectrometry. Using model cell lines for evaluation of TLR4 complex (TLR4/MD-2/CD14) stably transfected expression and verification of their transcription; these proteins could not be detected with conventional shotgun proteomics. Analysing recombinantly expressed and purified human TLR4 and MD-2 in human cell lines, both proteins were found to be glycosylated with high mannose structures. Compared to previously published reports of MD-2 glycosylation and the vendor's unwillingness to answer our questions regarding the method of recombinant protein purification, we cannot be sure these observed glycosylation profiles were representative of their mature forms in biologically relevant systems.

Using the recombinant proteins, we developed a targeted method for human TLR4 and MD-2 detection. The limits of detection for both TLR4 and MD-2 were 25fmol meaning that we needed at least 1500 copies of each protein per cell with a starting cell count of 1 million to detect these proteins. Using model cell lines for TLR4 activation, even with TLR4 activation, TLR4 was not detected however a single peptide corresponding to MD-2 was found. To understand our difficulty with TLR4 and MD-2 detection, all proteomics datasets published on a commonly used proteomics repository (PRIDE) were analysed. Manual analysis of these three datasets revealed TLR4 and MD-2 detection is rare in human proteomic datasets with less than 1% reporting detection of either protein. Manual

analysis of the spectral matches for the most promising of these datasets revealed only one out of five reported TLR4 peptides matches were of good quality.

9.2 FUTURE DIRECTIONS

In this study, we have improved the characterisation of glycans using LC retention time and by optimising mass spectrometric fragmentation spectra. Furthermore, we have developed internal standards and workflows to allow these improved methods to be widely applicable. Using software that is open-source and vendor-neutral is an essential part of this study, allowing our findings to be replicated and improved upon by other researchers in the field of glycomics, regardless of their analysis platform.

To follow up on our findings, there are two fundamental questions. The first, given that TLR4 has been investigated almost exclusively at the protein level by antibodies, is how specific previously utilised TLR4 antibodies are for their target. The second is how to best enrich for TLR4 given its low abundance in model cell line systems in order to allow for its detection, ultimately in *in vivo* systems. Both of these questions could be answered through an investigation of available TLR4 antibodies, validating their specificity and then using them as tools for enrichment prior to mass-spectrometric analysis. This approach has been utilised to quantitate the interacting partners of protein complexes with antibody enrichment and reduces the interference of artefacts from the a wide variety of protein pull-down methods[297]. Research into the TLR4 pathway could significantly benefit from verification of antibody specificity, detection and quantitation of TLR4 protein by mass spectrometry such that new experimental approaches for the study of the complex interactions in the pathway can be designed.

The application of these results for future development is currently limited by bioinformatic resources for glycomics data. From our knowledge, the diagnostic ion Skyline paper (Chapter 4) is the first glycomics study to provide associated raw data, diagnostic ion specificity and complete methods. Hopefully, as a result of MIRAGE and efforts with UniCarb-DB, more raw data can be available to further improve LC-MS-based glycomics using community-based knowledge. In my opinion, one of the biggest challenges in the field of glycomics is the heavy burden on researcher's knowledge regarding protein glycosylation due to the lack of automation of MS data interpretation. A combination of automated glycan isomer discrimination using diagnostic ions and retention time normalisation could serve as a useful step in establishing novice-friendly data analysis methods.

The study of *N*-glycan abundance changes following TLR4 activation provides a first glimpse into its effect on the cellular glycosylation pathway but only one subset was evaluated by analysing *N*-glycans with our PGC-LC-ESI-MS/MS method. For example, polysialic acid containing glycans, increasingly becoming of interest for understanding brain function, are not detected with our methods due to limitations with the organic solvent content used for glycan enrichment and

separation. Antibody-based methods appear to be the currently best available method for characterising the abundance of these structures. Additional gaps in our analysis of glycosylation changes following TLR4 activation is other forms of glycosylation: *O*-glycans, glycolipids and proteoglycans. While *O*-glycans were not covered by this study due to sample limitations, glycolipids and proteoglycans require alternative release and analysis methods.

Proteoglycans are a promising target for current and future research efforts, especially as research into the immune recognition of these degradation products have been further developed, identifying ligands responsible for self-associated molecular patterns consistent with chronic pain[298–300]. Only recently have methods made significant inroads towards characterisation of proteoglycans from limited biological samples[39, 49, 301] which, because of their increased relative abundance in the brain compared to the rest of the body[302], could prove promising for elucidating their role in the central nervous system. Analysis of these polysaccharides in the context of TLR4 activation could provide additional understanding of protein glycosylation changes following induction of neuropathic pain.

REFERENCES

1. Varki, A., Cummings, R.D., Aebi, M., Packer, N.H., Seeberger, P.H., Esko, J.D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J.J., Schnaar, R.L., Freeze, H.H., Marth, J.D., Bertozzi, C.R., Etzler, M.E., Frank, M., Vliegenthart, J.F.G., Lütteke, T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehisa, M., Toukach, P., Aoki-Kinoshita, K.F., Dell, A., Narimatsu, H., York, W., Taniguchi, N., Kornfeld, S.: Symbol nomenclature for graphical representations of glycans. Glycobiology. 25, 1323–1324 (2015).

2. Fenn, J.B.: Electrospray wings for molecular elephants (Nobel lecture). Angew. Chemie - Int. Ed. 42, 3871–3894 (2003).

3. Tanaka, K.: The origin of macromolecule ionization by laser irradiation (Nobel lecture). Angew. chemie Int. Ed. 197–217 (2003).

4. Wüthrich, K.: NMR studies of structure and function of biological macromolecules (Nobel Lecture). Angew. Chemie - Int. Ed. 42, 3340–3363 (2003).

5. Dehmelt, H.G., Paul, W.: The Nobel prize in physics: for the development of the ion trap technique. Stock. http//nobelprize. org/nobel_prizes/physics/laureates. (1989).

6. Ernst, R.R.: Nuclear Magnetic Resonance Fourier Transform Spectroscopy (Nobel Lecture) 1. Angew. chemie Int. Ed. 16, (1992).

7. Jonscher, K.R., Yates III, J.R.: The quadrupole ion trap mass spectrometer - a small solution to a big challenge. Anal. Biochem. 244, 1–15 (1997).

8. Jain, M., Koren, S., Miga, K.H., Quick, J., Rand, A.C., Sasani, T.A., Tyson, J.R., Beggs, A.D., Dilthey, A.T., Fiddes, I.T., Malla, S., Marriott, H., Nieto, T., O'Grady, J., Olsen, H.E., Pedersen, B.S., Rhie, A., Richardson, H., Quinlan, A.R., Snutch, T.P., Tee, L., Paten, B., Phillippy, A.M., Simpson, J.T., Loman, N.J., Loose, M.: Nanopore sequencing and assembly of a human genome with ultra-long reads. Nat. Biotechnol. 36, (2018).

9. Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., Williams, K.L., Hochstrasser, D.F.: From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Bio/Technology. 14, 61–65 (1996).

10. Edman, P., Begg, G.: A protein sequenator. Eur. J. Biochem. 1, 80–91 (1967).

11. Zhang, Y., Fonslow, B.R., Shan, B., Baek, M., Yates III, J.R.: Protein analysis by shotgun/bottom-up proteomics. Chem Rev.2. 113, 2343–2394 (2013).

12. Maccoss, M.J., Wu, C.C., Yates, J.R.: Probability-based validation of protein identifications

using a modified SEQUEST algorithm. Anal. Chem. 74, 5593–5599 (2002).

13. Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 20, 3551–3567 (1999).

14. Everest-Dass, A. V, Moh, E.S.X., Ashwood, C., Shathili, A.M.M., Packer, N.H.: Human disease glycomics: technology advances enabling protein glycosylation analysis – part 1. Expert Rev. Proteomics. 0, 14789450.2018.1421946 (2017).

15. Hayes, C.A., Karlsson, N.G., Struwe, W.B., Lisacek, F., Rudd, P.M., Packer, N.H., Campbell, M.P.: UniCarb-DB: A database resource for glycomic discovery. Bioinformatics. 27, 1343–1344 (2011).

16. Struwe, W.B., Agravat, S., Aoki-Kinoshita, K.F., Campbell, M.P., Costello, C.E., Dell, A., Feizi, T., Haslam, S.M., Karlsson, N.G., Khoo, K.H., Kolarich, D., Liu, Y., Mcbride, R., Novotny, M. V., Packer, N.H., Paulson, J.C., Rapp, E., Ranzinger, R., Rudd, P.M., Smith, D.F., Tiemeyer, M., Wells, L., York, W.S., Zaia, J., Kettner, C.: The minimum information required for a glycomics experiment (MIRAGE) project: sample preparation guidelines for reliable reporting of glycomics datasets. Glycobiology. 26, 907–910 (2016).

17. Kolarich, D., Rapp, E., Struwe, W.B., Haslam, S.M., Zaia, J., McBride, R., Agravat, S., Campbell, M.P., Kato, M., Ranzinger, R., Kettner, C., York, W.S.: The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data. Mol. Cell. Proteomics. 12, 991–5 (2013).

18. Alagesan, K., Khilji, S.K., Kolarich, D.: It is all about the solvent: on the importance of the mobile phase for ZIC-HILIC glycopeptide enrichment. Anal. Bioanal. Chem. 409, 529–538 (2017).

19. Pham, N.D., Pang, P.C., Krishnamurthy, S., Wands, A.M., Grassi, P., Dell, A., Haslam, S.M., Kohler, J.J.: Effects of altered sialic acid biosynthesis on N-linked glycan branching and cell surface interactions. J. Biol. Chem. 292, 9637–9651 (2017).

20. Hofmann, J., Stuckmann, A., Crispin, M., Harvey, D.J., Pagel, K., Struwe, W.B.: Identification of Lewis and Blood Group Carbohydrate Epitopes by Ion Mobility-Tandem-Mass Spectrometry Fingerprinting. Anal. Chem. 89, 2318–2325 (2017).

21. Möginger, U., Grunewald, S., Hennig, R., Kuo, C.-W., Schirmeister, F., Voth, H., Rapp, E., Khoo, K.-H., Seeberger, P.H., Simon, J.C., Kolarich, D.: Alterations of the Human Skin N- and O-Glycome in Basal Cell Carcinoma and Squamous Cell Carcinoma. Front. Oncol. 8, 70

(2018).

22. Ashwood, C., Lin, C.-H., Thaysen-Andersen, M., Packer, N.H.: Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. J. Am. Soc. Mass Spectrom. (2018).

23. Ashwood, C., Abrahams, J.L., Nevalainen, H., Packer, N.H.: Enhancing structural characterisation of glucuronidated O -linked glycans using negative mode ion trap-HCD mass spectrometry. Rapid Commun. Mass Spectrom. 31, 851–858 (2017).

24. Ramsby, M.L., Makowski, G.S., Khairallah, E. a: Differential detergent fractionation of isolated hepatocytes: biochemical, immunochemical and two-dimensional gel electrophoresis characterization of cytoskeletal and noncytoskeletal compartments. Electrophoresis. 15, 265–277 (1994).

25. Bradford, M.M.: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254 (1976).

26. Spiro, M.., Spiro, R.., Bhoyroo, V.D.: Studies on a thyroid enzyme involved in oligosaccharide transfer and the role of glucose in this reaction. J. Biol. Chem. 254, 7668–7674 (1979).

27. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, Mann, M.: In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856–2860 (2006).

28. Royle, L., Campbell, M.P., Radcliffe, C.M., White, D.M., Harvey, D.J., Abrahams, J.L., Kim, Y.G., Henry, G.W., Shadick, N.A., Weinblatt, M.E., Lee, D.M., Rudd, P.M., Dwek, R.A.: HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. Anal. Biochem. 376, 1–12 (2008).

29. Jensen, P.H., Karlsson, N.G., Kolarich, D., Packer, N.H.: Structural analysis of N- and O-glycans released from glycoproteins. Nat. Protoc. 7, 1299–1310 (2012).

30. Pilsoo, K., Mechref, Y., Novotny, M. V.: High-throughput solid-phase permethylation of glycans prior to mass spectrometry. Rapid Commun. Mass Spectrom. 22, 721–734 (2008).

31. Bern, M., Kil, Y.J., Becker, C.: Byonic: Advanced peptide and protein identification software. Curr. Protoc. Bioinforma. 1–14 (2012).

32. Weissgerber, T.L., Milic, N.M., Winham, S.J., Garovic, V.D.: Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. PLoS Biol. 13, 1–10 (2015).

33. Özdemir, A., Lin, J.-L., Wang, Y.S., Chen, C.-H.: A deeper look into sonic spray ionization. RSC Adv. 4, 61290–61297 (2014).

34. Harvey, D.J.: Fragmentation of negative ions from carbohydrates: Part 2. Fragmentation of high-mannose N-linked glycans. J. Am. Soc. Mass Spectrom. 16, 631–646 (2005).

35. Cech, N.B., Enke, C.G.: Practical implications of some recent studies in electrospray ionization fundamentals. Mass Spectrom. Rev. 20, 362–387 (2001).

36. Hsiao, C.-T., Wang, P.-W., Chang, H.-C., Chen, Y.-Y., Wang, S.-H., Chern, Y., Khoo, K.-H.: Advancing a high throughput glycotope-centric glycomics workflow based on nanoLC-MS2-product dependent-MS3 analysis of permethylated glycans. Mol. Cell. Proteomics. 2268–2280 (2017).

37. Schwartz, J.C., Senko, M.W., Syka, J.E.P.: A two-dimensional quadrupole ion trap mass spectrometer. J. Am. Soc. Mass Spectrom. 13, 659–669 (2002).

38. Bereman, M.S., Canterbury, J.D., Egertson, J.D., Horner, J., Remes, P.M., Schwartz, J.C., Zabrouskov, V., Maccoss, M.J.: Evaluation of front-end higher energy collision-induced dissociation on a bench-top dual pressure linear ion trap mass spectrometer for shotgun proteomics. Anal. Chem. 84, 1533–1539 (2012).

39. Hogan, J.D., Klein, J.A., Wu, J., Chopra, P., Boons, G.-J., Carvalho, L., Lin, C., Zaia, J.: Software for peak finding and elemental composition assignment for glycosaminoglycan tandem mass spectra. Mol. Cell. Proteomics. 1, 1–33 (2018).

40. Charlwood, J., Langridge, J., Camilleri, P.: Structural characterisation of N-linked glycan mixtures by precursor ion scanning and tandem mass spectrometric analysis. Rapid Commun. Mass Spectrom. 13, 1522–1530 (1999).

41. Pabst, M., Bondili, J.S., Stadlmann, J., Mach, L., Altmann, F.: Mass + retention time = structure: A strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. Anal. Chem. 79, 5051–5057 (2007).

42. Vreeker, G.C.M., Wuhrer, M.: Reversed-phase separation methods for glycan analysis. Anal. Bioanal. Chem. 409, 359–378 (2017).

43. Zaia, J.: Mass spectrometry of oligosaccharides. Mass Spectrom. Rev. 23, 161–227 (2004).

44. Moh, E.S.X., Thaysen-Andersen, M., Packer, N.H.: Relative versus absolute quantitation in disease glycomics. Proteomics - Clin. Appl. 9, 368–382 (2015).

45. Ruhaak, L.R., Zauner, G., Huhn, C., Bruggink, C., Deelder, a. M., Wuhrer, M.: Glycan labeling strategies and their use in identification and quantification. Anal. Bioanal. Chem. 397,

3457–3481 (2010).

46. Guillarme, D., Schappler, J., Rudaz, S., Veuthey, J.-L.: Coupling ultra-high-pressure liquid chromatography with mass spectrometry. TrAC Trends Anal. Chem. 29, 15–27 (2010).

47. Hasehira, K., Tateno, H., Onuma, Y., Ito, Y., Asashima, M., Hirabayashi, J.: Structural and Quantitative Evidence for Dynamic Glycome Shift on Production of Induced Pluripotent Stem Cells. Mol. Cell. Proteomics. 11, 1913–1923 (2012).

48. Wuhrer, M., Deelder, A.M., van der Burgt, Y.E..: Mass spectrometric glycan rearrangements. Mass Spectrom. Rev. 30, 664–680 (2011).

49. Klein, J.A., Meng, L., Zaia, J.: Deep sequencing of complex proteoglycans: a novel strategy for high coverage and site- specific identification of glycosaminoglycan-linked peptides. (2018).

50. Fenaille, F., Barbier Saint-Hilaire, P., Rousseau, K., Junot, C.: Data acquisition workflows in liquid chromatography coupled to high resolution mass spectrometry-based metabolomics: Where do we stand? J. Chromatogr. A. 1526, 1–12 (2017).

51. Adamski, J., Suhre, K.: Metabolomics platforms for genome wide association studies-linking the genome to the metabolome, (2013).

52. Geyer, H., Schmitt, S., Wuhrer, M., Geyer, R.: Structural analysis of glycoconjugates by on-target enzymatic digestion and MALDI-TOF-MS. Anal. Chem. 71, 476–482 (1999).

53. Wu, L., Han, D.K.: Overcoming the dynamic range problem in mass spectrometry-based shotgun proteomics. Expert Rev. Proteomics. 3, 611–619 (2006).

54. Nefedov, A. V., Mitra, I., Brasier, A.R., Sadygov, R.G.: Examining troughs in the mass distribution of all theoretically possible tryptic peptides. J. Proteome Res. 10, 4150–4157 (2011).

55. Akune, Y., Lin, C.H., Abrahams, J.L., Zhang, J., Packer, N.H., Aoki-Kinoshita, K.F., Campbell, M.P.: Comprehensive analysis of the N-glycan biosynthetic pathway using bioinformatics to generate UniCorn: A theoretical N-glycan structure database. Carbohydr. Res. 431, 56–63 (2016).

56. Yu, G., Zhang, Y., Zhang, Z., Song, L., Wang, P., Chai, W.: Effect and limitation of excess ammonium on the release of O-glycans in reducing forms from glycoproteins under mild alkaline conditions for glycomic and functional analysis. Anal. Chem. 82, 9534–9542 (2010).

57. Lauc, G., Essafi, A., Huffman, J.E., Hayward, C., Kneevi, A., Kattla, J.J., Polacek, O., Gornik, O.,

Vitart, V., Abrahams, J.L., Pui, M., Novokmet, M., Redi, I., Campbell, S., Wild, S.H., Boroveki, F., Wang, W., Koli, I., Zgaga, L., Gyllensten, U., Wilson, J.F., Wright, A.F., Hastie, N.D., Campbell, H., Rudd, P.M., Rudan, I.: Genomics meets glycomics-the first gwas study of human N-glycome identifies HNF1A as a master regulator of plasma protein fucosylation. PLoS Genet. 6, 1–14 (2010).

58. Gillet, L.C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., Aebersold, R.: Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. Mol. Cell. Proteomics. 11, O111.016717 (2012).

59. Geiger, T., Cox, J., Mann, M.: Proteomics on an Orbitrap Benchtop Mass Spectrometer Using All-ion Fragmentation. Mol. Cell. Proteomics. 9, 2252–2261 (2010).

60. Reinhold, V.N., Sheeley, D.M.: Detailed characterization of carbohydrate linkage and sequence in an ion trap mass spectrometer: Glycosphingolipids. Anal. Biochem. 259, 28–33 (1998).

61. Harvey, D.J., Abrahams, J.L.: Fragmentation and ion mobility properties of negative ions from N-linked carbohydrates: Part 7. Reduced glycans. Rapid Commun. Mass Spectrom. 30, 627–634 (2016).

62. Fenn, L.S., McLean, J.A.: Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry. Phys. Chem. Chem. Phys. 13, 2196–2205 (2011).

63. Ruhaak, L.R., Lebrilla, C.B.: Applications of Multiple Reaction Monitoring to Clinical Glycomics. Chromatographia. 78, 335–342 (2015).

64. Smith, J., Mittermayr, S., Váradi, C., Bones, J.: Quantitative glycomics using liquid phase separations coupled to mass spectrometry. Analyst. 142, 700–720 (2017).

65. Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., Coon, J.J.: Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. Mol. Cell. Proteomics. 11, 1475–1488 (2012).

66. Rockwood, A.L., Busman, M., Udseth, H.R., Smith, R.D.: Thermally induced dissociation of ions from electrospray mass spectrometry. Rapid Commun. Mass Spectrom. 5, 582–585 (1991).

67. Chataigné, G., Couderc, F., Poinsot, V.: Polysaccharides analysis of sinorhizobial capside by on-line anion exchange chromatography with pulsed amperometric detection and mass spectrometry coupling. J. Chromatogr. A. 1185, 241–250 (2008).

68. Yamashita, M., Fenn, J.B.: Negative ion production with the electrospray ion source. J. Phys. Chem. 88, 4671–4675 (1984).

69. Mechref, Y., Pilsoo, K., Novotny, M. V.: Differentiating structural isomers of sialylated glycans by matrix-assisted laser desorption/ionization time-of-flight/time-of-flight tandem mass spectrometry. Rapid Commun. Mass Spectrom. 20, 1381–1389 (2006).

70. Ritamo, I., Räbinä, J., Natunen, S., Valmu, L.: Nanoscale reversed-phase liquid chromatography-mass spectrometry of permethylated N-glycans. Anal. Bioanal. Chem. 405, 2469–2480 (2013).

71. Costello, C.E., Contado-Miller, J.M., Cipollo, J.F.: A Glycomics Platform for the Analysis of Permethylated Oligosaccharide Alditols. J. Am. Soc. Mass Spectrom. 18, 1799–1812 (2007).

72. Cerdán-Calero, M., Sendra, J.M., Sentandreu, E.: Gas chromatography coupled to mass spectrometry analysis of volatiles, sugars, organic acids and aminoacids in Valencia Late orange juice and reliability of the Automated Mass Spectral Deconvolution and Identification System for their automatic identifica. J. Chromatogr. A. 1241, 84–95 (2012).

73. Nesvizhskii, A.I., Roos, F.F., Grossmann, J., Vogelzang, M., Eddes, J.S., Gruissem, W., Baginsky, S., Aebersold, R.: Dynamic Spectrum Quality Assessment and Iterative Computational Analysis of Shotgun Proteomic Data. Mol. Cell. Proteomics. 5, 652–670 (2006).

74. Cooper, C.A., Gasteiger, E., Packer, N.H.: GlycoMod--a software tool for determining glycosylation compositions from mass spectrometric data. Proteomics. 1, 340–349 (2001).

75. Zhang, Z., Marshall, A.G.: A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. J. Am. Soc. Mass Spectrom. 9, 225–233 (1998).

76. Armbruster, D.A., Pry, T.: Limit of blank, limit of detection and limit of quantitation. Clin. Biochem. Rev. 29 Suppl 1, S49-52 (2008).

77. Leymarie, N., Zaia, J.: Effective Use of Mass Spectrometry for Glycan and Glycopeptide Structural Analysis. (2012).

78. Douglas, D.J.: Mechanism of the collision-induced dissociation of polyatomic ions studied by triple quadrupole mass spectrometry. J. Phys. Chem. 86, 185–191 (1982).

79. Zhao, C., Xie, B., Chan, S.Y., Costello, C.E., O'Connor, P.B.: Collisionally Activated

Dissociation and Electron Capture Dissociation Provide Complementary Structural Information for Branched Permethylated Oligosaccharides. J. Am. Soc. Mass Spectrom. 19, 138–150 (2008).

80. Goeringer, D.E., McLuckey, S.A.: Kinetics of collision-induced dissociation in the paul trap: A first-order model. Rapid Commun. Mass Spectrom. 10, 328–334 (1996).

81. McLuckey, S.A., Goeringer, D.E.: Slow heating methods in tandem mass spectrometry. J. Mass Spectrom. 32, 461–474 (1997).

82. Löner, C., Blackstock, W., Gunaratne, J.: Enhanced performance of pulsed q collision induced dissociation-based peptide identification on a dual-pressure linear ion trap. J. Am. Soc. Mass Spectrom. 23, 186–189 (2012).

83. Tang, H., Mechref, Y., Novotny, M. V.: Automated interpretation of MS/MS spectra of oligosaccharides. Bioinformatics. 21, 431–439 (2005).

84. Ranzinger, R., Weatherly, D., Arpinar, S., Khan, S., Porterfield, M., Tiemeyer, M., S. York, W.: GRITS Toolbox - A software system for the archival, processing and interpretation of glycomics data. In: Annual Meeting of the Society-for-Glycobiology on Glycobiology. p. 1275 (2015).

85. Yu, C.Y., Mayampurath, A., Hu, Y., Zhou, S., Mechref, Y., Tang, H.: Automated annotation and quantification of glycans using liquid chromatography-mass spectrometry. Bioinformatics. 29, 1706–1707 (2013).

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

87. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., Haslam, S.M.: GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. J. Proteome Res. 7, 1650–1659 (2008).

88. Everest-Dass, A. V., Abrahams, J.L., Kolarich, D., Packer, N.H., Campbell, M.P.: Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. J. Am. Soc. Mass Spectrom. 24, 895–906 (2013).

89. Webb, I.K., Chen, T.C., Danielson, W.F., Ibrahim, Y.M., Tang, K., Anderson, G.A., Smith, R.D.: Implementation of dipolar resonant excitation for collision induced dissociation with ion mobility/time-of-flight MS. J. Am. Soc. Mass Spectrom. 25, 563–571 (2014).

90. MacLean, B., Tomazela, D.M., Abbatiello, S.E., Zhang, S., Whiteaker, J.R., Paulovich, A.G., Carr, S.A., MacCoss, M.J.: Effect of collision energy optimization on the measurement of

peptides by selected reaction monitoring (SRM) mass spectrometry. Anal. Chem. 82, 10116–10124 (2010).

91. Harvey, D.J., Royle, L., Radcliffe, C.M., Rudd, P.M., Dwek, R.A.: Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. Anal. Biochem. 376, 44–60 (2008).

92. Przybylski, C., Bonnet, V.: Discrimination of cyclic and linear oligosaccharides by tandem mass spectrometry using collision-induced dissociation (CID), pulsed-Q-dissociation (PQD) and the higherenergy C-trap dissociation modes. Rapid Commun. Mass Spectrom. 27, 75–87 (2013).

93. Sharma, V., Eckels, J., Taylor, G.K., Shulman, N.J., Stergachis, A.B., Joyner, S.A., Yan, P., Whiteaker, J.R., Halusa, G.N., Schilling, B., Gibson, B.W., Colangelo, C.M., Paulovich, A.G., Carr, S.A., Jaffe, J.D., Maccoss, M.J., Maclean, B.: Panorama: A targeted proteomics knowledge base. J. Proteome Res. 13, 4205–4210 (2014).

94. Harrison, M.J., Nouwens, A.S., Jardine, D.R., Zachara, N.E., Gooley, A.A., Nevalainen, H., Packer, N.H.: Modified glycosylation of cellobiohydrolase I from a high cellulase-producing mutant strain of Trichoderma reesei. Eur. J. Biochem. 256, 119–127 (1998).

95. Guile, G.R., Rudd, P.M., Wing, D.R., Prime, S.B., Dwek, R.A.: A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. Anal. Biochem. 240, 210–226 (1996).

96. Rudd, P.M., Rudan, I., Wright, A.F.: High-throughput glycome analysis is set to join high-throughput genomics. J. Proteome Res. 8, 1105 (2009).

97. Melmer, M., Stangler, T., Premstaller, A., Lindner, W.: Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis. J. Chromatogr. A. 1218, 118–123 (2011).

98. Pabst, M., Altmann, F.: Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides. Anal. Chem. 80, 7534–7542 (2008).

99. Aldredge, D., An, H.J., Tang, N., Waddell, K., Lebrilla, C.B.: Annotation of a serum N-Glycan library for rapid identification of structures. J. Proteome Res. 11, 1958–1968 (2012).

100. Abrahams, J.L., Campbell, M.P., Packer, N.H.: Building a PGC-LC-MS N-glycan retention library and elution mapping resource. Glycoconj. J. 35, 15–29 (2018).

101. Luo, Q., Rejtar, T., Wu, S.-L., Karger, B.L.: Hydrophilic interaction 10 μm I.D. porous layer

open tubular columns for ultratrace glycan analysis by liquid chromatography-mass spectrometry. J Chromatogr. A. 1216, 1223–1231 (2009).

102. Ahn, J., Bones, J., Yu, Y.Q., Rudd, P.M., Gilar, M.: Separation of 2-aminobenzamide labeled glycans using hydrophilic interaction chromatography columns packed with 1.7 μm sorbent. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 878, 403–408 (2010).

103. Stockmann, H., Adamczyk, B., Hayes, J., Rudd, P.M.: Automated, high-throughput IgG-antibody glycoprofiling platform. Anal. Chem. 85, 8841–8849 (2013).

104. Bundle, D.R., Jennings, H.J., Kenny, C.P.: Studies on the group-specific polysaccharide of Neisseria meningitidis serogroup X and an improved procedure for its isolation. J. Biol. Chem. 249, 4797–4801 (1974).

105. Campbell, M.P., Royle, L., Radcliffe, C.M., Dwek, R.A., Rudd, P.M.: GlycoBase and autoGU: Tools for HPLC-based glycan analysis. Bioinformatics. 24, 1214–1216 (2008).

106. Schneider, P., Ferguson, M.A.J.: Microscale Analysis of Glycosylphosphatidylinositol Structures. Methods Enzymol. 250, 614–630 (1995).

107. Masuko, T., Minami, A., Iwasaki, N., Majima, T., Nishimura, S.I., Lee, Y.C.: Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. Anal. Biochem. 339, 69–72 (2005).

108. Jin, C., Kenny, D.T., Skoog, E.C., Padra, M., Adamczyk, B., Vitizeva, V., Thorell, A., Venkatakrishnan, V., Lindén, S.K., Karlsson, N.G.: Structural diversity of human gastric mucin glycans. Mol. Cell. Proteomics. mcp.M117.067983 (2017).

109. MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., MacCoss, M.J.: Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 26, 966–968 (2010).

110. Thompson, J.W., Pratt, B., Horowitz-gelb, M., Dubois, L.G., John-williams, L.S., Astarita, G., Moseley, M.A., Maccoss, M., Maclean, B.: Skyline for small molecules : a flexible tool for cross-platform LC-MS / MS method creation and data analysis for metabolomics . 3.

111. Kessner, D., Chambers, M., Burke, R., Agus, D., Mallick, P.: ProteoWizard: Open source software for rapid proteomics tools development. Bioinformatics. 24, 2534–2536 (2008).

112. Tiemeyer, M., Aoki, K., Paulson, J., Cummings, R.D., York, W.S., Karlsson, N.G., Lisacek, F., Packer, N.H., Campbell, M.P., Aoki, N.P., Fujita, A., Matsubara, M., Shinmachi, D., Tsuchiya, S., Yamada, I., Pierce, M., Ranzinger, R., Narimatsu, H., Aoki-Kinoshita, K.F.: GlyTouCan: An

accessible glycan structure repository. Glycobiology. 27, 915–919 (2017).

113. Frewen, B., MacCoss, M.J.: Using BiblioSpec for Creating and Searching Tandem MS Peptide Libraries. Curr. Protoc. Bioinforma. 1–12 (2007).

114. Forgács, E., Cserháti, T.: Relationship between retention characteristics and physicochemical parameters of solutes on porous graphitized carbon column. J. Pharm. Biomed. Anal. 18, 505–510 (1998).

115. Zhou, S., Hu, Y., Mechref, Y.: High-temperature LC-MS/MS of permethylated glycans derived from glycoproteins. Electrophoresis. 37, 1506–1513 (2016).

116. Mizuno, H., Ueda, K., Kobayashi, Y., Tsuyama, N., Todoroki, K., Min, J.Z., Toyo'oka, T.: The great importance of normalization of LC–MS data for highly-accurate non-targeted metabolomics. Biomed. Chromatogr. 31, 1–7 (2017).

117. Elbein, A.: Glycosylation Inhibitors for N-Linked Glycoproteins. In: Methods in Enzymology. pp. 661–709 (1987).

118. Zaia, J.: Mass Spectrometry and the Emerging Field of Glycomics, (2008).

119. Shubhakar, A., Reiding, K.R., Gardner, R.A., Spencer, D.I.R., Fernandes, D.L., Wuhrer, M.: High-Throughput Analysis and Automation for Glycomics Studies. Chromatographia. 78, 321–333 (2014).

120. Čmelík, R., Štikarovská, M., Chmelík, J.: Different behavior of dextrans in positive-ion and negative-ion mass spectrometry. J. Mass Spectrom. 39, 1467–1473 (2004).

121. Toprak, U.H., Gillet, L.C., Maiolica, A., Navarro, P., Leitner, A., Aebersold, R.: Conserved Peptide Fragmentation as a Benchmarking Tool for Mass Spectrometers and a Discriminating Feature for Targeted Proteomics. Mol. Cell. Proteomics. 13, 2056–2071 (2014).

122. Tang, Y., Wei, J., Costello, C.E., Lin, C.: Characterization of Isomeric Glycans by Reversed Phase Liquid Chromatography-Electronic Excitation Dissociation Tandem Mass Spectrometry. J. Am. Soc. Mass Spectrom. 1–13 (2018).

123. Manz, C., Pagel, K.: Glycan analysis by ion mobility-mass spectrometry and gas-phase spectroscopy. Curr. Opin. Chem. Biol. 42, 16–24 (2018).

124. Hayase, T., Lee, Y.C., Rice, K.G., Dziegielewska, K.M., Mark, K., Thomas, R.: Comparison of N-Glycosides of Fetuins from Different Species and Human α2-HS-Glycoprotein. Biochemistry. 31, 4915–4921 (1992).

125. Spik, G., Coddeville, B., Mazurier, J., Bourne, Y., Cambillaut, C., Montreuil, J.: Primary and Three-Dimensional Structure of Lactotransferrin (Lactoferrin) Glycans. In: Lactoferrin Structure and Function. pp. 21–33 (1994).

126. Loke, I., Østergaard, O., Heegaard, N.H.H., Packer, N.H., Thaysen-Andersen, M.: Paucimannose-Rich N -glycosylation of Spatiotemporally Regulated Human Neutrophil Elastase Modulates Its Immune Functions. Mol. Cell. Proteomics. 16, 1507–1527 (2017).

127. Pučić, M., Knežević, A., Vidič, J., Adamczyk, B., Novokmet, M., Polašek, O., Gornik, O., Šupraha-Goreta, S., Wormald, M.R., Redžić, I., Campbell, H., Wright, A., Hastie, N.D., Wilson, J.F., Rudan, I., Wuhrer, M., Rudd, P.M., Josić, D., Lauc, G.: High Throughput Isolation and Glycosylation Analysis of IgG–Variability and Heritability of the IgG Glycome in Three Isolated Human Populations. Mol. Cell. Proteomics. 10, M111.010090 (2011).

128. Mattu, T.S., Pleass, R.J., Willis, A.C., Kilian, M., Wormald, M.R., Lellouch, A.C., Rudd, P.M., Woof, J.M., Dwek, R.A.: The glycosylation and structure of human serum IgA1, Fab, and Fc regions and the role of N-glycosylation on Fcα receptor interactions. J. Biol. Chem. 273, 2260–2272 (1998).

129. Hui, J.P.M., Lanthier, P., White, T.C., McHugh, S.G., Yaguchi, M., Roy, R., Thibault, P.: Characterization of cellobiohydrolase I (Cel7A) glycoforms from extracts of Trichoderma reesei using capillary isoelectric focusing and electrospray mass spectrometry. J. Chromatogr. B Biomed. Sci. Appl. 752, 349–368 (2001).

130. Thomsson, K.A., Karlsson, H., Hansson, G.C.: Sequencing of sulfated oligosaccharides from mucins by liquid chromatography and electrospray ionization tandem mass spectrometry. Anal. Chem. 72, 4543–9 (2000).

131. Bianchi, M.E.: DAMPs, PAMPs and alarmins: all we need to know about danger. J. Leukoc. Biol. 81, 1–5 (2007).

132. Lehnardt, S., Lachance, C., Patrizi, S., Lefebvre, S., Follett, P.L., Jensen, F.E., Rosenberg, P. a, Volpe, J.J., Vartanian, T.: The toll-like receptor TLR4 is necessary for lipopolysaccharide-induced oligodendrocyte injury in the CNS. J. Neurosci. 22, 2478–2486 (2002).

133. Murphy, T.J., Paterson, H.M., Kriynovich, S., Zang, Y., Kurt-jones, E. a, Mannick, J. a, Lederer, J. a: Linking the " two-hit " response following injury to enhanced TLR4 reactivity Abstract : Severe injury can initiate an exagger- ated systemic inflammatory response and multiple organ failure (MOF) if a subsequent immune stim-. J. Leukoc. Biol. (2004).

134. Gargiulo, S., Gamba, P., Testa, G., Rossin, D., Biasi, F., Poli, G., Leonarduzzi, G.: Relation between TLR4/NF-κB signaling pathway activation by 27-hydroxycholesterol and 4-

hydroxynonenal, and atherosclerotic plaque instability. Aging Cell. 569–581 (2015).

135. Schulthess, F.T., Paroni, F., Sauter, N.S., Shu, L., Ribaux, P., Haataja, L., Strieter, R.M., Oberholzer, J., King, C.C., Maedler, K.: CXCL10 Impairs β Cell Function and Viability in Diabetes through TLR4 Signaling. Cell Metab. 9, 125–139 (2009).

136. Kim, F., Pham, M., Luttrell, I., Bannerman, D.D., Tupper, J., Thaler, J., Hawn, T.R., Raines, E.W., Schwartz, M.W.: Toll-like receptor-4 mediates vascular inflammation and insulin resistance in diet-induced obesity. Circ. Res. 100, 1589–1596 (2007).

137. Fernandez-Lizarbe, S., Pascual, M., Guerri, C.: Critical role of TLR4 response in the activation of microglia induced by ethanol. J. Immunol. 183, 4733–4744 (2009).

138. Lewis, S.S., Hutchinson, M.R., Zhang, Y., Hund, D.K., Maier, S.F., Rice, K.C., Watkins, L.R.: Glucuronic acid and the ethanol metabolite ethyl-glucuronide cause toll-like receptor 4 activation and enhanced pain. Brain. Behav. Immun. 30, 24–32 (2013).

139. Oyama, J.I., Blais, C., Liu, X., Pu, M., Kobzik, L., Kelly, R. a., Bourcier, T.: Reduced Myocardial Ischemia-Reperfusion Injury in Toll-Like Receptor 4-Deficient Mice. Circulation. 109, 784–789 (2004).

140. Berglund, L., Björling, E., Oksvold, P., Fagerberg, L., Asplund, A., Al-Khalili Szigyarto, C., Persson, A., Ottosson, J., Wernérus, H., Nilsson, P., Lundberg, E., Sivertsson, Å., Navani, S., Wester, K., Kampf, C., Hober, S., Pontén, F., Uhlén, M.: A Genecentric Human Protein Atlas for Expression Profiles Based on Antibodies. Mol. Cell. Proteomics. 7, 2019–2027 (2008).

141. Hutchinson, M.R., Zhang, Y., Shridhar, M., Evans, J.H., Buchanan, M.M., Zhao, T.X., Slivka, P.F., Coats, B.D., Rezvani, N., Wieseler, J., Hughes, T.S., Landgraf, K.E., Chan, S., Fong, S., Phipps, S., Falke, J.J., Leinwand, L. a., Maier, S.F., Yin, H., Rice, K.C., Watkins, L.R.: Evidence that opioids may have toll-like receptor 4 and MD-2 effects. Brain. Behav. Immun. 24, 83–95 (2010).

142. Tramullas, M., Finger, B.C., Moloney, R.D., Golubeva, A. V., Moloney, G., Dinan, T.G., Cryan, J.F.: Toll-like receptor 4 regulates chronic stress-induced visceral pain in mice. Biol. Psychiatry. 76, 340–348 (2014).

143. Hua, F., Ma, J., Ha, T., Kelley, J.L., Kao, R.L., Schweitzer, J.B., Kalbfleisch, J.H., Williams, D.L., Li, C.: Differential roles of TLR2 and TLR4 in acute focal cerebral ischemia/reperfusion injury in mice. Brain Res. 1262, 100–108 (2009).

144. Pascual, O., Ben Achour, S., Rostaing, P., Triller, A., Bessis, A.: Microglia activation triggers astrocyte-mediated modulation of excitatory neurotransmission. Proc. Natl. Acad. Sci. 109, E197–E205 (2012).

145. Valanne, S., Wang, J.-H., Ramet, M.: The Drosophila Toll Signaling Pathway. J. Immunol. 186, 649–656 (2011).

146. Hetru, C., Hoffmann, J.A.: NF-KB in the Immune Response of Drosophila. Cold Spring Harb. Perspect. Biol. 1, a000232–a000232 (2009).

147. Sultzer, B.M.: Genetic control of host responses to endotoxin. Infect. Immun. 5, 107–113 (1972).

148. Poltorak, A., Smirnova, I., He, X., Liu, M.Y., Van Huffel, C., Birdwell, D., Alejos, E., Silva, M., Du, X., Thompson, P., Chan, E.K.L., Ledesma, J., Roe, B., Clifton, S., Vogel, S.N., Beutler, B.: Genetic and physical mapping of the Lps locus: Identification of the toll-4 receptor as a candidate gene in the critical region. Blood Cells, Mol. Dis. 24, 340–355 (1998).

149. Milligan, E.D., Watkins, L.R.: Pathological and protective roles of glia in chronic pain. 10, (2009).

150. Fitzgerald, K. a, Rowe, D.C., Barnes, B.J., Caffrey, D.R., Visintin, A., Latz, E., Monks, B., Pitha, P.M., Golenbock, D.T.: LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF. J. Exp. Med. 198, 1043–1055 (2003).

151. Medzhitov, R., Preston-Hurlburt, P., Kopp, E., Stadlen, a, Chen, C., Ghosh, S., Janeway, C. a: MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. Mol. Cell. 2, 253–258 (1998).

152. Boukhvalova, M.S., Prince, G.A., Soroush, L., Harrigan, D.C., Vogel, S.N., Blanco, J.C.G.: The TLR4 agonist, monophosphoryl lipid A, attenuates the cytokine storm associated with respiratory syncytial virus vaccine-enhanced disease. Vaccine. 24, 5027–5035 (2006).

153. Park, B.S., Lee, J.-O.: Recognition of lipopolysaccharide pattern by TLR4 complexes. Exp. Mol. Med. 45, e66 (2013).

154. Zhao, M., Zhou, A., Xu, L., Zhang, X.: The role of TLR4-mediated PTEN/PI3K/AKT/NF-κB signaling pathway in neuroinflammation in hippocampal neurons. Neuroscience. 269, 93–101 (2014).

155. Watkins, L.R., Milligan, E.D., Maier, S.F.: Glial activation: A driving force for pathological pain. Trends Neurosci. 24, 450–455 (2001).

156. Grace, P.M., Hutchinson, M.R., Maier, S.F., Watkins, L.R.: Pathological pain and the neuroimmune interface. Nat. Rev. Immunol. 14, 217–31 (2014).

157. Kim, H.M., Park, B.S., Kim, J.I., Kim, S.E., Lee, J., Oh, S.C., Enkhbayar, P., Matsushima, N., Lee, H., Yoo, O.J., Lee, J.O.: Crystal Structure of the TLR4-MD-2 Complex with Bound Endotoxin

Antagonist Eritoran. Cell. 130, 906–917 (2007).

158. Matsunaga, N., Tsuchimori, N., Matsumoto, T., Ii, M.: TAK-242 (resatorvid), a small-molecule inhibitor of Toll-like receptor (TLR) 4 signaling, binds selectively to TLR4 and interferes with interactions between TLR4 and its adaptor molecules. Mol. Pharmacol. 79, 34–41 (2011).

159. Neal, M.D., Jia, H., Eyer, B., Good, M., Guerriero, C.J., Sodhi, C.P., Afrazi, A., Jr, T.P., Ma, C., Branca, M., Ozolek, J., Brodsky, J.L., Wipf, P., Hackam, D.J.: Discovery and Validation of a New Class of Small Molecule Toll-Like Receptor 4 (TLR4) Inhibitors. 8, 1–10 (2013).

160. Bettoni, I., Comelli, F., Rossini, C., Granucci, F., Giagnoni, G., Peri, F., Costa, B.: Glial TLR4 receptor as new target to treat neuropathic pain: Efficacy of a new receptor antagonist in a model of peripheral nerve injury in mice. Glia. 56, 1312–1319 (2008).

161. Hutchinson, M.R., Zhang, Y., Brown, K., Coats, B.D., Shridhar, M., Sholar, P.W., Patel, S.J., Crysdale, N.Y., Harrison, J. a., Maier, S.F., Rice, K.C., Watkins, L.R.: Non-stereoselective reversal of neuropathic pain by naloxone and naltrexone: Involvement of toll-like receptor 4 (TLR4). Eur. J. Neurosci. 28, 20–29 (2008).

162. Jiang, Q., Akashi, S., Miyake, K., Petty, H.R.: Cutting Edge: Lipopolysaccharide Induces Physical Proximity Between CD14 and Toll-Like Receptor 4 (TLR4) Prior to Nuclear Translocation of NF- B. J. Immunol. 165, 3541–3544 (2000).

163. Takahashi, Y., Isuzugawa, K., Murase, Y., Imai, M., Yamamoto, S.: Up-Regulation of NOD1 and NOD2 through TLR4 and TNF- α in LPS-treated Murine Macrophages. (2006).

164. Sheedy, F.J., Palsson-Mcdermott, E., Hennessy, E.J., Martin, C., O'Leary, J.J., Ruan, Q., Johnson, D.S., Chen, Y., O'Neill, L.A.J.: Negative regulation of TLR4 via targeting of the proinflammatory tumor suppressor PDCD4 by the microRNA miR-21. Nat. Immunol. 11, 141–147 (2010).

165. Hill, M., Tanguy-Royer, S., Royer, P., Chauveau, C., Asghar, K., Tesson, L., Lavainne, F., Rémy, S., Brion, R., Hubert, F.X., Heslan, M., Rimbert, M., Berthelot, L., Moffett, J.R., Josien, R., Grégoire, M., Anegon, I.: IDO expands human CD4+CD25high regulatory T cells by promoting maturation of LPS-treated dendritic cells. Eur. J. Immunol. 37, 3054–3062 (2007).

166. Neal, M.D., Leaphart, C., Levy, R., Prince, J., Billiar, T.R., Watkins, S., Li, J., Cetin, S., Ford, H., Schreiber, A., Hackam, D.J.: Enterocyte TLR4 mediates phagocytosis and translocation of bacteria across the intestinal barrier. J. Immunol. 176, 3070–3079 (2006).

167. Wu, F.X., Bian, J.J., Miao, X.R., Huang, S.D., Xu, X.W., Gong, D.J., Sun, Y.M., Lu, Z.J., Yu, W.F.:

Intrathecal siRNA against toll-like receptor 4 reduces nociception in a rat model of neuropathic pain. Int. J. Med. Sci. 7, 251–259 (2010).

168. Austin, P.J., Wu, A., Moalem-Taylor, G.: Chronic Constriction of the Sciatic Nerve and Pain Hypersensitivity Testing in Rats. J. Vis. Exp. 1–6 (2012).

169. Bennett, G.J., Xie, Y.K.: A peripheral mononeuropathy in rat that produces d... [Pain. 1988] - PubMed result. Pain. 33, 87–107 (1988).

170. Watkins, L.R., Maier, S.F., Goehler, L.E.: Immune activation: the role of pro-inflammatory cytokines in inflammation, illness responses and pathological pain states. Pain. 63, 289–302 (1995).

171. Jarvis, M.F.: The neural-glial purinergic receptor ensemble in chronic pain states. Trends Neurosci. 33, 48–57 (2010).

172. Apkarian, A.V., Bushnell, M.C., Treede, R.D., Zubieta, J.K.: Human brain mechanisms of pain perception and regulation in health and disease. Eur. J. Pain. 9, 463–484 (2005).

173. Pinho, S.S., Reis, C.A.: Glycosylation in cancer: Mechanisms and clinical implications. Nat. Rev. Cancer. 15, 540–555 (2015).

174. Freeze, H.H., Eklund, E.A., Ng, B.G., Patterson, M.C.: Neurological Aspects of Human Glycosylation Disorders. Annu. Rev. Neurosci. 38, 105–125 (2015).

175. Wallace, R.: Extending Tlusty's rate distortion index theorem method to the glycome: Do even "low level" biochemical phenomena require sophisticated cognitive paradigms? BioSystems. 107, 145–152 (2012).

176. MacAuley, M.S., Crocker, P.R., Paulson, J.C.: Siglec-mediated regulation of immune cell function in disease. Nat. Rev. Immunol. 14, 653–666 (2014).

177. Schnaar, R.L.: Glycans and glycan-binding proteins in immune regulation: A concise introduction to glycobiology for the allergist. J. Allergy Clin. Immunol. 135, 609–615 (2015).

178. Maverakis, E., Kim, K., Shimoda, M., Gershwin, M.E., Patel, F., Wilken, R., Raychaudhuri, S., Ruhaak, L.R., Lebrilla, C.B.: Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: A critical review. J. Autoimmun. 57, 1–13 (2015).

179. Teghanemt, A., Widstrom, R.L., Gioannini, T.L., Weiss, J.P.: Isolation of monomeric and dimeric secreted MD-2: EndotoxinsCD14 and Toll-like receptor 4 ectodomain selectively react with the monomeric form of secreted MD-2. J. Biol. Chem. 283, 21881–21889 (2008).

180. Visintin, a, Mazzoni, a, Spitzer, J. a, Segal, D.M.: Secreted MD-2 is a large polymeric protein that efficiently confers lipopolysaccharide sensitivity to Toll-like receptor 4. Proc. Natl. Acad. Sci. U. S. A. 98, 12156–12161 (2001).

181. Flo, T.H., Ryan, L., Latz, E., Takeuchi, O., Monks, B.G., Lien, E., Halaas, O., Akira, S., Skjåk-Bræk, G., Golenbock, D.T., Espevik, T.: Involvement of Toll-like receptor (TLR) 2 and TLR4 in cell activation by mannuronic acid polymers. J. Biol. Chem. 277, 35489–35495 (2002).

182. Shimizu, T., Hirano, H., Shimizu, S., Kishioka, C., Sakakura, Y., Majima, Y.: Differential Properties of Mucous Glycoproteins in Rat Nasal Epithelium. Am J Respir Crit Care Med. 164, 1077-0182 (2001).

183. Iijima, J., Kobayashi, S., Kitazume, S., Kizuka, Y., Fujinawa, R., Korekane, H., Shibata, T., Saitoh, S.I., Akashi-Takamura, S., Miyake, K., Miyoshi, E., Taniguchi, N.: Core fucose is critical for CD14-dependent Toll-like receptor 4 signaling. Glycobiology. 27, 1006–1015 (2017).

184. Kim, S.M., Kim, B.Y., Eo, S.K., Kim, C.D., Kim, K.: 27-Hydroxycholesterol up-regulates CD14 and predisposes monocytic cells to superproduction of CCL2 in response to lipopolysaccharide. Biochim. Biophys. Acta - Mol. Basis Dis. 1852, 442–450 (2015).

185. Tsuda, M., Inoue, K., Salter, M.W.: Neuropathic pain and spinal microglia: A big problem from molecules in “small” glia. Trends Neurosci. 28, 101–107 (2005).

186. Li, Y., Dorsi, M.J., Meyer, R.A., Belzberg, A.J.: Mechanical hyperalgesia after an L5 spinal nerve lesion in the rat is not dependent on input from injured nerve fibers. Pain. 85, 493–502 (2000).

187. Moisset, X., Bouhassira, D.: Brain imaging of neuropathic pain. Neuroimage. 37, 80–88 (2007).

188. Abrahams, J.L., Packer, N.H., Campbell, M.P.: Relative quantitation of multi-antennary N-glycan classes: combining PGC-LC-ESI-MS with exoglycosidase digestion. Analyst. (2015).

189. Wang, Y., He, H., Li, D., Zhu, W., Duan, K., Le, Y., Liao, Y., Ou, Y.: The role of the TLR4 signaling pathway in cognitive deficits following surgery in aged rats. Mol. Med. Rep. 7, 1137–1142 (2013).

190. Loyd, D.R., Murphy, A.Z.: The role of the periaqueductal gray in the modulation of pain in males and females: Are the anatomy and physiology really that different? Neural Plast. 2009, (2009).

191. Jensen, M.P., Karoly, P., Braver, S.: The measurement of clinical pain intensity: a

comparison of six methods. Pain. 27, 117–126 (1986).

192. Dahmen, A.C., Fergen, M.T., Laurini, C., Schmitz, B., Loke, I., Thaysen-Andersen, M., Diestel, S.: Paucimannosidic glycoepitopes are functionally involved in proliferation of neural progenitor cells in the subventricular zone. Glycobiology. 25, 869–880 (2015).

193. Stevens, C.W., Kajander, K.C., Bennett, G.J., Seybold, V.S.: Bilateral and differential changes in spinal mu, delta and kappa opioid binding in rats with a painful, unilateral neuropathy. Pain. 46, 315–326 (1991).

194. Honore, P., Rogers, S.D., Schwei, M.J., Salak-Johnson, J.L., Luger, N.M., Sabino, M.C., Clohisy, D.R., Mantyh, P.W.: Murine models of inflammatory, neuropathic and cancer pain each generates a unique set of neurochemical changes in the spinal cord and sensory neurons. Neuroscience. 98, 585–598 (2000).

195. Chen, Y.J., Wing, D.R., Guile, G.R., Dwek, R. a, Harvey, D.J., Zamze, S.: Neutral N-glycans in adult rat brain tissue--complete characterisation reveals fucosylated hybrid and complex structures. Eur. J. Biochem. 251, 691–703 (1998).

196. Zamze, S., Harvey, D.J., Chen, Y.J., Guile, G.R., Dwek, R.A., Wing, D.R.: Sialylated N-glycans in adult rat brain tissue - A widespread distribution of disialylated antennae in complex and hybrid structures. Eur. J. Biochem. 258, 243–270 (1998).

197. Schwalbe, R.A., Corey, M.J., Cartwright, T.A.: Novel Kv3 glycoforms differentially expressed in adult mammalian brain contain sialylated N - glycans. Biochem. Cell Biol. 86, 21–30 (2008).

198. Nakakita, S.I., Natsuka, S., Okamoto, J., Ikenaka, K., Hase, S.: Alteration of brain type N-glycans in neurological mutant mouse brain. J. Biochem. 138, 277–283 (2005).

199. Fawcett, J.W., Asher, R.A.: The glial scar and central nervous system repair. Brain Res. Bull. 49, 377–391 (1999).

200. Schnaar, R.L., Gerardy-Schahn, R., Hildebrandt, H.: Sialic acids in the brain: gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. Physiol. Rev. 94, 461–518 (2014).

201. Ye, X., Johann, D.J., Hakami, R.M., Xiao, Z., Meng, Z., Ulrich, R.G., Issaq, H.J., Veenstra, T.D., Blonder, J.: Optimization of protein solubilization for the analysis of the CD14 human monocyte membrane proteome using LC-MS/MS. J. Proteomics. 73, 112–122 (2009).

202. Yan, J.X., Wilkins, M.R., Ou, K., Gooley, A.A., Williams, K.L., Sanchez, J.C., Golaz, O., Pasquali, C., Hochstrasser, D.F.: Large-scale amino-acid analysis for proteome studies. J.

Chromatogr. A. 736, 291–302 (1996).

203. Peng, J., Gygi, S.P.: Proteomics: The move to mixtures. J. Mass Spectrom. 36, 1083–1091 (2001).

204. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B.: Quantitative mass spectrometry in proteomics: A critical review. Anal. Bioanal. Chem. 389, 1017–1031 (2007).

205. Aggarwal, K., Choe, L.H., Lee, K.H.: Shotgun proteomics using the iTRAQ isobaric tags. Briefings Funct. Genomics Proteomics. 5, 112–120 (2006).

206. Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Hamon, C.: Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. 75, 1895–1904 (2003).

207. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., Mann, M.: Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. Mol. Cell. Proteomics. 1, 376–386 (2002).

208. Vaudel, M., Burkhart, J.M., Radau, S., Zahedi, R.P., Martens, L., Sickmann, A.: Integral quantification accuracy estimation for reporter ion-based quantitative proteomics (iQuARI). J. Proteome Res. 11, 5072–5080 (2012).

209. Ishihama, Y.: Proteomic LC-MS systems using nanoscale liquid chromatography with tandem mass spectrometry. J. Chromatogr. A. 1067, 73–83 (2005).

210. Schilling, B., Rardin, M.J., MacLean, B.X., Zawadzka, a. M., Frewen, B.E., Cusack, M.P., Sorensen, D.J., Bereman, M.S., Jing, E., Wu, C.C., Verdin, E., Kahn, C.R., MacCoss, M.J., Gibson, B.W.: Platform-independent and Label-free Quantitation of Proteomic Data Using MS1 Extracted Ion Chromatograms in Skyline: APPLICATION TO PROTEIN ACETYLATION AND PHOSPHORYLATION. Mol. Cell. Proteomics. 11, 202–214 (2012).

211. Wuhr, M., Haas, W., McAlister, G.C., Peshkin, L., Rad, R., Kirschner, M., Gygi, S.P.: Accurate Multiplexed Proteomics at the MS2 Level Using the Complement Reporter Ion Cluster. Anal Chem. 84, 9214–9221 (2012).

212. Ting, L., Rad, R., Gygi, S.P., Haas, W.: MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat. Methods. 8, 937–940 (2011).

213. Pappin, D.J.C., Hojrup, P., Bleasby, A.J.: Rapid identification of proteins by peptide-mass fingerprinting. Curr. Biol. 3, 327–332 (1993).

214. Hunt, D.F., Yates, J.R., Shabanowitz, J., Winston, S., Hauer, C.R.: Protein sequencing by

tandem mass spectrometry. Proc. Natl. Acad. Sci. 83, 6233–6237 (1986).

215. Payne, A.H., Chelf, J.H., Glish, G.L.: C-terminal peptide sequencing using acetylated peptides with MSn in a quadrupole ion trap. Analyst. 125, 635–640 (2000).

216. Michalski, A., Cox, J., Mann, M.: More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. J. Proteome Res. 10, 1785–1793 (2011).

217. Xie, H., Gilar, M., Gebler, J.C.: Characterization of protein impurities and site-specific modifications using peptide mapping with liquid chromatography and data independent acquisition mass spectrometry. Anal. Chem. 81, 5699–5708 (2009).

218. Picotti, P., Aebersold, R.: Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions. Nat. Methods. 9, 555–566 (2012).

219. Lange, V., Picotti, P., Domon, B., Aebersold, R.: Selected reaction monitoring for quantitative proteomics: a tutorial. Mol. Syst. Biol. 4, 222 (2008).

220. Worboys, J.D., Sinclair, J., Yuan, Y., Jørgensen, C.: Systematic evaluation of quantotypic peptides for targeted analysis of the human kinome. Nat. Methods. 11, 1041–1044 (2014).

221. Shan, S.W., Do, C.W., Lam, T.C., Kong, R.P.W., Li, K.K., Chun, K.M., Stamer, W.D., To, C.H.: New Insight of Common Regulatory Pathways in Human Trabecular Meshwork Cells in Response to Dexamethasone and Prednisolone Using an Integrated Quantitative Proteomics: SWATH and MRM-HR Mass Spectrometry. J. Proteome Res. 16, 3753–3765 (2017).

222. Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., Weissman, J.S.: Global analysis of protein expression in yeast. Nature. 425, 737–741 (2003).

223. Shi, T., Song, E., Nie, S., Rodland, K.D., Liu, T., Qian, W., Smith, D.: Advances in targeted proteomics and applications to biomedical research. Proteomics. 16, 2160–2182 (2016).

224. Woo, J., Han, D., Wang, J.I., Park, J., Kim, H., Kim, Y.: Quantitative Proteomics Reveals Temporal Proteomic Changes in Signaling Pathways during BV2 Mouse Microglial Cell Activation. J. Proteome Res. 16, 3419–3432 (2017).

225. Martens, L.: Bringing proteomics into the clinic: The need for the field to finally take itself seriously. Proteomics - Clin. Appl. 7, 388–391 (2013).

226. Martens, L., Hermjakob, H., Jones, P., Adamsk, M., Taylor, C., States, D., Gevaert, K., Vandekerckhove, J., Apweiler, R.: PRIDE: The proteomics identifications database.

Proteomics. 5, 3537–3545 (2005).

227. Perez-Riverol, Y., Xu, Q.-W., Wang, R., Uszkoreit, J., Griss, J., Sanchez, A., Reisinger, F., Csordas, A., Ternent, T., Del-Toro, N., Dianes, J.A., Eisenacher, M., Hermjakob, H., Vizcaíno, J.A.: PRIDE Inspector Toolsuite: Moving Toward a Universal Visualization Tool for Proteomics Data Standard Formats and Quality Assessment of ProteomeXchange Datasets. Mol. Cell. Proteomics. 15, 305–317 (2016).

228. Park, B.S., Song, D.H., Kim, H.M., Choi, B.-S., Lee, H., Lee, J.-O.: The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. Nature. 458, 1191–1195 (2009).

229. Andersen, D.C., Krummen, L.: Recombinant protein expression for therapeutic applications. Curr. Opin. Biotechnol. 13, 117–123 (2002).

230. Da Silva Correia, J., Ulevitch, R.J.: MD-2 and TLR4 N-linked glycosylations are important for a functional lipopolysaccharide receptor. J. Biol. Chem. 277, 1845–1854 (2002).

231. Cario, E., Podolsky, D.K.: Differential alteration in intestinal epithelial cell expression of toll-like receptor 3 (TLR3) and TLR4 in inflammatory bowel disease. Infect. Immun. 68, 7010–7017 (2000).

232. Nattermann, J., Du, X., Wei, Y., Shevchenko, D., Beutler, B.: Endotoxin-mimetic effect of antibodies against Toll-like receptor 4. J Endotoxin Res. 6, 257–264 (2000).

233. Arbour, N.C., Lorenz, E., Schutte, B.C., Zabner, J., Kline, J.N., Jones, M., Frees, K., Watt, J.L., Schwartz, D. a: TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. Nat. Genet. 25, 187–191 (2000).

234. Ji, Y., Liu, J., Wang, Z., Liu, N.: Angiotensin II induces inflammatory response partly via toll-like receptor 4-dependent signaling pathway in vascular smooth muscle cells. Cell Physiol Biochem. 23, 265–276 (2009).

235. Hietaranta, A., Mustonen, H., Puolakkainen, P., Haapiainen, R., Kemppainen, E.: Proinflammatory effects of pancreatic elastase are mediated through TLR4 and NF-κB. Biochem. Biophys. Res. Commun. 323, 192–196 (2004).

236. Taylor, K.R., Trowbridge, J.M., Rudisill, J. a., Termeer, C.C., Simon, J.C., Gallo, R.L.: Hyaluronan Fragments Stimulate Endothelial Recognition of Injury through TLR4. J. Biol. Chem. 279, 17079–17084 (2004).

237. Takenaka, R., Yokota, K., Ayada, K., Mizuno, M., Zhao, Y., Fujinami, Y., Lin, S.N., Toyokawa, T., Okada, H., Shiratori, Y., Oguma, K.: Helicobacter pylori heat-shock protein 60 induces

inflammatory responses through the Toll-like receptor-triggered pathway in cultured human gastric epithelial cells. Microbiology. 150, 3913–3922 (2004).

238. Kawahara, T., Kuwano, Y., Teshima-Kondo, S., Kawai, T., Nikawa, T., Kishi, K., Rokutan, K.: Toll-like receptor 4 regulates gastric pit cell responses to Helicobacter pylori infection. J Med Invest. 48, 190–197 (2001).

239. Tomita, M., Ohkubo, R., Hayashi, M.: Lipopolysaccharide Transport System across Colonic Epithelial Cells in Normal and Infective Rat. Drug Metab. Pharmacokinet. 19, 33–40 (2004).

240. Fazeli, A., Bruce, C., Anumba, D.O.: Characterization of Toll-like receptors in the female reproductive tract in humans. Hum. Reprod. 20, 1372–1378 (2005).

241. Sugawara, Y., Uehara, A., Fujimoto, Y., Kusumoto, S., Fukase, K., Shibata, K., Sugawara, S., Sasano, T., Takada, H.: Toll-like receptors, NOD1, and NOD2 in oral epithelial cells. J. Dent. Res. 65, 524–529 (2006).

242. Bomfim, G.F., Dos Santos, R.A., Oliveira, M.A., Giachini, F.R., Akamine, E.H., Tostes, R.C., Fortes, Z.B., Webb, R.C., Carvalho, M.H.: Toll like receptor 4 contributes to blood pressure regulation and vascular contraction in spontaneously hypertensive rat. Clin. Sci. 122, 535–543 (2012).

243. Guillott, L., Medjane, S., Le-Barillec, K., Balloy, V., Danel, C., Chignard, M., Si-Tahar, M.: Response of human pulmonary epithelial cells to lipopolysaccharide involves toll-like receptor 4 (TLR4)-dependent signaling pathways: Evidence for an intracellular compartmentalization of TLR4. J. Biol. Chem. 279, 2712–2718 (2004).

244. Da Silva Correia, J., Soldau, K., Christen, U., Tobias, P.S., Ulevitch, R.J.: Lipopolysaccharide is in close proximity to each of the proteins in its membrane receptor complex. Transfer from CD14 to TLR4 and MD-2. J. Biol. Chem. 276, 21129–21135 (2001).

245. Bordeaux, J., Welsh, A.W., Agarwal, S., Killiam, E., Baquero, M.T., Hanna, J.A., Anagnostou, V.K., Rimm, D.L.: Antibody validation. Biotechniques. 48, 197–209 (2010).

246. Uhlen, M., Bandrowski, A., Carr, S., Edwards, A., Ellenberg, J., Lundberg, E., Rimm, D.L., Rodriguez, H., Hiltke, T., Snyder, M., Yamamoto, T.: A proposal for validation of antibodies. Nat. Methods. 13, 823–827 (2016).

247. Pal, D., Dasgupta, S., Kundu, R., Maitra, S., Das, G., Mukhopadhyay, S., Ray, S., Majumdar, S.S., Bhattacharya, S.: Fetuin-A acts as an endogenous ligand of TLR4 to promote lipid-induced insulin resistance. Nat. Med. 18, 1279–1285 (2012).

248. Huang, Z., Zhao, C., Chen, Y., Cowell, J.A., Wei, G., Kultti, A., Huang, L., Thompson, C.B., Rosengren, S., Frost, G.I., Shepard, H.M.: Recombinant Human Hyaluronidase PH20 Does Not Stimulate an Acute Inflammatory Response and Inhibits Lipopolysaccharide-Induced Neutrophil Recruitment in the Air Pouch Model of Inflammation. J. Immunol. 192, 5285–5295 (2014).

249. Ishida, T., Mizushina, Y., Yagi, S., Irino, Y., Nishiumi, S., Miki, I., Kondo, Y., Mizuno, S., Yoshida, H., Azuma, T., Yoshida, M.: Inhibitory Effects of Glycyrrhetinic Acid on DNA Polymerase and Inflammatory Activities. Evidence-Based Complement. Altern. Med. 2012, 1–9 (2012).

250. Park, M.J., Ryu, H.S., Kim, J.S., Lee, H.K., Kang, J.S., Yun, J., Kim, S.Y., Lee, M.K., Hong, J.T., Kim, Y., Han, S.B.: Platycodon grandiflorum polysaccharide induces dendritic cell maturation via TLR4 signaling. Food Chem. Toxicol. 72, 212–220 (2014).

251. Barberá, A., Lorenzo, N., Garrido, G., Mazola, Y., Falcón, V., Torres, A.M., Hernández, M.I., Hernández, M.V., Margry, B., De Groot, A.M., Van Roon, J., Van Der Zee, R., Broere, F., Van Eden, W., Padrón, G., Del Carmen Domínguez, M.: APL-1, an altered peptide ligand derived from human heat-shock protein 60, selectively induces apoptosis in activated CD4 + CD25 + T cells from peripheral blood of rheumatoid arthritis patients. Int. Immunopharmacol. 17, 1075–1083 (2013).

252. Mazurek, J., Ignatowicz, L., Kallenius, G., Svenson, S.B., Pawlowski, A., Hamasur, B.: Divergent effects of mycobacterial cell wall glycolipids on maturation and function of human monocyte-derived dendritic cells. PLoS One. 7, 1–13 (2012).

253. Wang, X., Zhang, Y., Peng, Y., Hutchinson, M.R., Rice, K.C., Yin, H., Watkins, L.R.: Pharmacological characterization of the opioid inactive isomers (+)-naltrexone and (+)-naloxone as antagonists of toll-like receptor 4. Br. J. Pharmacol. 173, 856–869 (2016).

254. Merline, R., Moreth, K., Beckmann, J., Nastase, M. V., Zeng-Brouwers, J., Tralhao, J.G., Lemarchand, P., Pfeilschifter, J., Schaefer, R.M., Iozzo, R. V., Schaefer, L.: Signaling by the Matrix Proteoglycan Decorin Controls Inflammation and Cancer Through PDCD4 and MicroRNA-21. Sci. Signal. 4, ra75-ra75 (2011).

255. Needham, B.D., Carroll, S.M., Giles, D.K., Georgiou, G., Whiteley, M., Trent, M.S.: Modulating the innate immune response by combinatorial engineering of endotoxin. Proc. Natl. Acad. Sci. 110, 1464–1469 (2013).

256. Mamat, U., Wilke, K., Bramhill, D., Schromm, A.B., Lindner, B., Kohl, T.A., Corchero, J.L., Villaverde, A., Schaffer, L., Head, S.R., Souvignier, C., Meredith, T.C., Woodard, R.W.: Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins.

Microb. Cell Fact. 14, 1–15 (2015).

257. Malpede, B.M., Lin, D.H., Tolia, N.H.: Molecular basis for sialic acid-dependent receptor recognition by the Plasmodium falciparum invasion protein erythrocyte-binding antigen-140/baebl. J. Biol. Chem. 288, 12406–12415 (2013).

258. Lee, Y.H., Jeon, S.H., Kim, S.H., Kim, C., Lee, S.J., Koh, D., Lim, Y., Ha, K., Shin, S.Y.: A new synthetic chalcone derivative, 2-hydroxy-3',5,5'-trimethoxychalcone (DK-139), suppresses the toll-like receptor 4-mediated inflammatory response through inhibition of the Akt/NF-KB pathway in BV2 microglial cells. Exp. Mol. Med. 44, 369–377 (2012).

259. Jessen, D.L., Osei-Owusu, P., Toosky, M., Roughead, W., Bradley, D.S., Nilles, M.L.: Type III secretion needle proteins induce cell signaling and cytokine secretion via toll-like receptors. Infect. Immun. 82, 2300–2309 (2014).

260. Zhang, B., Yin, Y., Lai, R.C., Tan, S.S., Choo, A.B.H., Lim, S.K.: Mesenchymal Stem Cells Secrete Immunologically Active Exosomes. Stem Cells Dev. 23, 1233–1244 (2014).

261. Dadaglio, G., Fayolle, C., Zhang, X., Ryffel, B., Oberkampf, M., Felix, T., Hervas-Stubbs, S., Osicka, R., Sebo, P., Ladant, D., Leclerc, C.: Antigen Targeting to CD11b $^{+}$ Dendritic Cells in Association with TLR4/TRIF Signaling Promotes Strong CD8 $^{+}$ T Cell Responses. J. Immunol. 193, 1787–1798 (2014).

262. Stevens, C.W., Aravind, S., Das, S., Davis, R.L.: Pharmacological characterization of LPS and opioid interactions at the toll-like receptor 4. Br. J. Pharmacol. 168, 1421–1429 (2013).

263. Yamaba, S., Yamada, S., Kajikawa, T., Awata, T., Sakashita, H., Tsushima, K., Fujihara, C., Yanagita, M., Murakami, S.: PLAP-1/asporin regulates TLR2- and TLR4-induced inflammatory responses. J. Dent. Res. 94, 1706–1714 (2015).

264. Fiorotto, R., Scirpo, R., Trauner, M., Fabris, L., Hoque, R., Spirli, C., Strazzabosco, M.: Loss of CFTR affects biliary epithelium innate immunity and causes TLR4NF-κB-mediated inflammatory response in mice. Gastroenterology. 141, 1498–1508.e5 (2011).

265. Gilmore, T.D.: Introduction to NF-κB: Players, pathways, perspectives. Oncogene. 25, 6680–6684 (2006).

266. Xu, L.G., Wang, Y.Y., Han, K.J., Li, L.Y., Zhai, Z., Shu, H.B.: VISA is an adapter protein required for virus-triggered IFN-β signaling. Mol. Cell. 19, 727–740 (2005).

267. Ben Haij, N., Leghmari, K., Planès, R., Thieblemont, N., Bahraoui, E.: HIV-1 Tat protein binds to TLR4-MD2 and signals to induce TNF-α and IL-10. Retrovirology. 10, 1–12 (2013).

268. Thomson, S.J.P., Goh, F.G., Banks, H., Krausgruber, T., Kotenko, S. V, Foxwell, B.M.J., Udalova, I.A.: The role of transposable elements in the regulation of IFN-lambda1 gene expression. Proc. Natl. Acad. Sci. U. S. A. 106, 11564–9 (2009).

269. Shembade, N., Parvatiyar, K., Harhaj, N.S., Harhaj, E.W.: The ubiquitin-editing enzyme A20 requires RNF11 to downregulate NF-κB signalling. EMBO J. 28, 513–522 (2009).

270. Krausgruber, T., Saliba, D., Ryzhakov, G., Lanfrancotti, A., Blazek, K., Udalova, I.A.: IRF5 is required for late-phase TNF secretion by human dendritic cells. Blood. 115, 4421–4430 (2010).

271. Into, T., Inomata, M., Nakashima, M., Shibata, K. -i., Hacker, H., Matsushita, K.: Regulation of MyD88-Dependent Signaling Events by S Nitrosylation Retards Toll-Like Receptor Signal Transduction and Initiation of Acute-Phase Immune Responses. Mol. Cell. Biol. 28, 1338–1347 (2008).

272. Necela, B.M., Su, W., Thompson, E.A.: Toll-like receptor 4 mediates cross-talk between peroxisome proliferator-activated receptor and nuclear factor-KB in macrophages. Immunology. 125, 344–358 (2008).

273. Hornung, V., Rothenfusser, S., Britsch, S., Krug, A., Jahrsdorfer, B., Giese, T., Endres, S., Hartmann, G.: Quantitative Expression of Toll-Like Receptor 1-10 mRNA in Cellular Subsets of Human Peripheral Blood Mononuclear Cells and Sensitivity to CpG Oligodeoxynucleotides. J. Immunol. 168, 4531–4537 (2002).

274. Han, J., Kim, H.J., Lee, S.C., Hong, S., Park, K., Jeon, Y.H., Kim, D., Cheong, H.K., Kim, H.S.: Structure-based rational design of a Toll-like receptor 4 (TLR4) decoy receptor with high binding affinity for a target protein. PLoS One. 7, (2012).

275. Thaysen-Andersen, M., Thøgersen, I.B., Nielsen, H.J., Lademann, U., Brünner, N., Enghild, J.J., Højrup, P.: Rapid and individual-specific glycoprofiling of the low abundance N-glycosylated protein tissue inhibitor of metalloproteinases-1. Mol. Cell. Proteomics. 6, 638–647 (2007).

276. Wu, J.X., Song, X., Pascovici, D., Zaw, T., Care, N., Krisp, C., Molloy, M.P.: SWATH Mass Spectrometry Performance Using Extended Peptide MS/MS Assay Libraries. Mol. Cell. Proteomics. 15, 2501–2514 (2016).

277. Dittrich, J., Adam, M., Maas, H., Hecht, M., Reinicke, M., Ruhaak, R., Cobbaert, C., Engel, C., Wirkner, K., Löffler, M., Thiery, J., Ceglarek, U.: Targeted on-line SPE-LC-MS/MS Assay for the quantitation of twelve apolipoproteins from human blood. Proteomics. 1700279 (2017).

278. Guo, X., Trudgian, D.C., Lemoff, A., Yadavalli, S., Mirzaei, H.: Confetti: A Multiprotease Map of the *HeLa* Proteome for Comprehensive Proteomics. Mol. Cell. Proteomics. 13, 1573–1584 (2014).

279. Olsen, J. V., Ong, S.-E., Mann, M.: Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. Mol. Cell. Proteomics. 3, 608–614 (2004).

280. Nagai, Y., Akashi, S., Nagafuku, M., Ogata, M., Iwakura, Y., Akira, S., Kitamura, T., Kosugi, A., Kimoto, M., Miyake, K.: Essential role of MD-2 in LPS responsiveness and TLR4 distribution. Nat. Immunol. 3, 667–672 (2002).

281. Ohnishi, T., Muroi, M., Tanamoto, K.-I.: MD-2 is necessary for the toll-like receptor 4 protein to undergo glycosylation essential for its translocation to the cell surface. Clin. Diagn. Lab. Immunol. 10, 405–410 (2003).

282. Mirzaei, H., Regnier, F.: Enhancing electrospray ionization efficiency of peptides by derivatization. Anal. Chem. 78, 4175–4183 (2006).

283. Malvandi, A.M., Mehrzad, J., Moghaddam, M.S.: Gene expression quantification of toll like receptors 2, 4 and co-molecules in Human Glioblastoma cell line (U87-MG): Toward a new in vitro model of inflammation. Iran. J. Basic Med. Sci. 14, 428–435 (2011).

284. Gong, Z., Calkins, G., Cheng, E., Krause, D., Niklason, L.E.: Influence of Culture Medium on Smooth Muscle Cell Differentiation from Human Bone Marrow–Derived Mesenchymal Stem Cells. Tissue Eng. Part A. 15, 319–330 (2009).

285. Gangloff, M., Gay, N.J.: MD-2: The Toll "gatekeeper" in endotoxin signalling. Trends Biochem. Sci. 29, 294–300 (2004).

286. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T.L.: NCBI BLAST: a better web interface. Nucleic Acids Res. 36, 5–9 (2008).

287. Li, X., Tran, K.M., Aziz, K.E., Sorokin, A. V., Chen, J., Wang, W.: Defining the Protein-Protein Interaction Network of the Human Protein Tyrosine Phosphatase Family. Mol. Cell. Proteomics. 15, 3030–3044 (2016).

288. Lee, J., McKinney, K.Q., Pavlopoulos, A.J., Han, M.H., Kim, S.H., Kim, H.J., Hwang, S.: Exosomal proteome analysis of cerebrospinal fluid detects biosignatures of neuromyelitis optica and multiple sclerosis. Clin. Chim. Acta. 462, 118–126 (2016).

289. Kollipara, L., Buchkremer, S., Weis, J., Brauers, E., Hoss, M., Rütten, S., Caviedes, P., Zahedi, R.P., Roos, A.: Proteome Profiling and Ultrastructural Characterization of the Human RCMH Cell Line: Myoblastic Properties and Suitability for Myopathological Studies. J.

Proteome Res. 15, 945–955 (2016).

290. Wang, X., Zhang, B., Wren, J.: CustomProDB: An R package to generate customized protein databases from RNA-Seq data for proteomics search. Bioinformatics. 29, 3235–3237 (2013).

291. Zhou, R., Kuang, Y., Zhou, J., Du, X., Li, J., Shi, J., Haburcak, R., Xu, B.: Nanonets collect cancer secretome from pericellular space. PLoS One. 11, 1–18 (2016).

292. Kil, Y.J., Bern, M., Crowell, K., Kletter, D., Bern, N., Tang, W., Carlson, E., Becker, C.: Towards a Comprehensive Bioinformatic Analysis of the NIST Reference mAb. (2015).

293. Lu, Y.-C., Yeh, W.-C., Ohashi, P.S.: LPS/TLR4 signal transduction pathway. Cytokine. 42, 145–151 (2008).

294. Guzzo, C., Ayer, A., Basta, S., Banfield, B.W., Gee, K.: IL-27 Enhances LPS-Induced Proinflammatory Cytokine Production via Upregulation of TLR4 Expression and Signaling in Human Monocytes. J. Immunol. 188, 864–873 (2012).

295. Wiśniewski, J.R., Hein, M.Y., Cox, J., Mann, M.: A "Proteomic Ruler" for Protein Copy Number and Concentration Estimation without Spike-in Standards. Mol. Cell. Proteomics. 13, 3497–3506 (2014).

296. Anderson, L., Hunter, C.L.: Quantitative Mass Spectrometric Multiple Reaction Monitoring Assays for Major Plasma Proteins. Mol. Cell. Proteomics. 5, 573–588 (2006).

297. Trinkle-Mulcahy, L., Boulon, S., Lam, Y.W., Urcia, R., Boisvert, F.M., Vandermoere, F., Morrice, N.A., Swift, S., Rothbauer, U., Leonhardt, H., Lamond, A.: Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. J. Cell Biol. 183, 223–239 (2008).

298. Brunn, G.J., Bungum, M.K., Johnson, G.B., Platt, J.L.: Conditional signaling by Toll-like receptor 4. FASEB J. 19, 872–874 (2005).

299. Moreth, K., Iozzo, R. V., Schaefer, L.: Small leucine-rich proteoglycans orchestrate receptor crosstalk during inflammation. Cell Cycle. 11, 2084–2091 (2012).

300. O'Callaghan, P., Li, J.-P., Lannfelt, L., Lindahl, U., Zhang, X.: Microglial heparan sulfate proteoglycans facilitate the cluster-of-differentiation 14 (CD14)/Toll-like receptor 4 (TLR4)-dependent inflammatory response. J. Biol. Chem. 290, jbc.M114.634337 (2015).

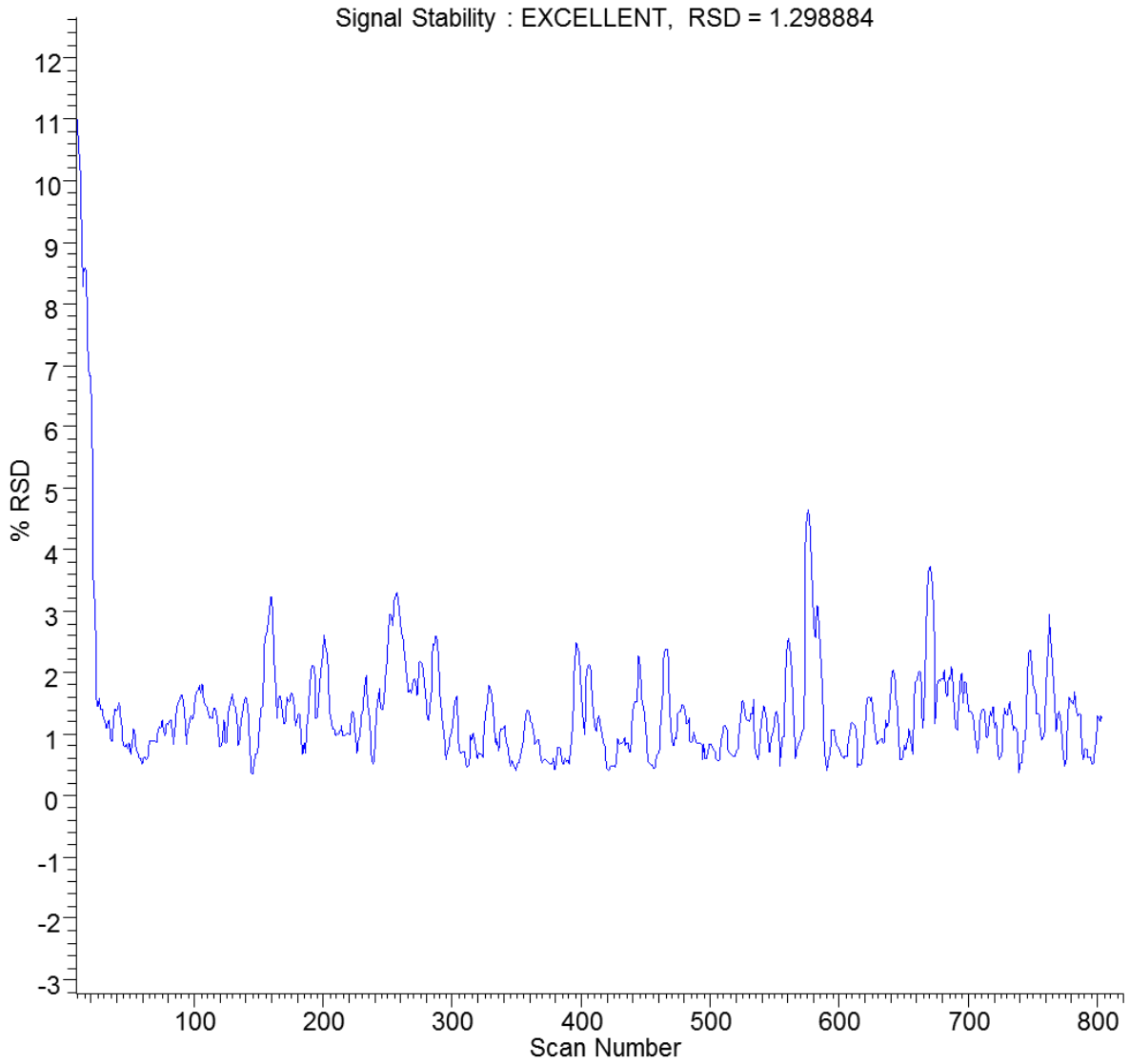
301. Turiák, L., Tóth, G., Ozohanics, O., Révész, Á., Ács, A., Vékey, K., Zaia, J., Drahos, L.: Sensitive method for glycosaminoglycan analysis of tissue sections. J. Chromatogr. A. 1544, 41–48 (2018).

302. Zimmermann, D.R., Dours-Zimmermann, M.T.: Extracellular matrix of the central nervous system: From neglect to challenge. Histochem. Cell Biol. 130, 635–653 (2008).

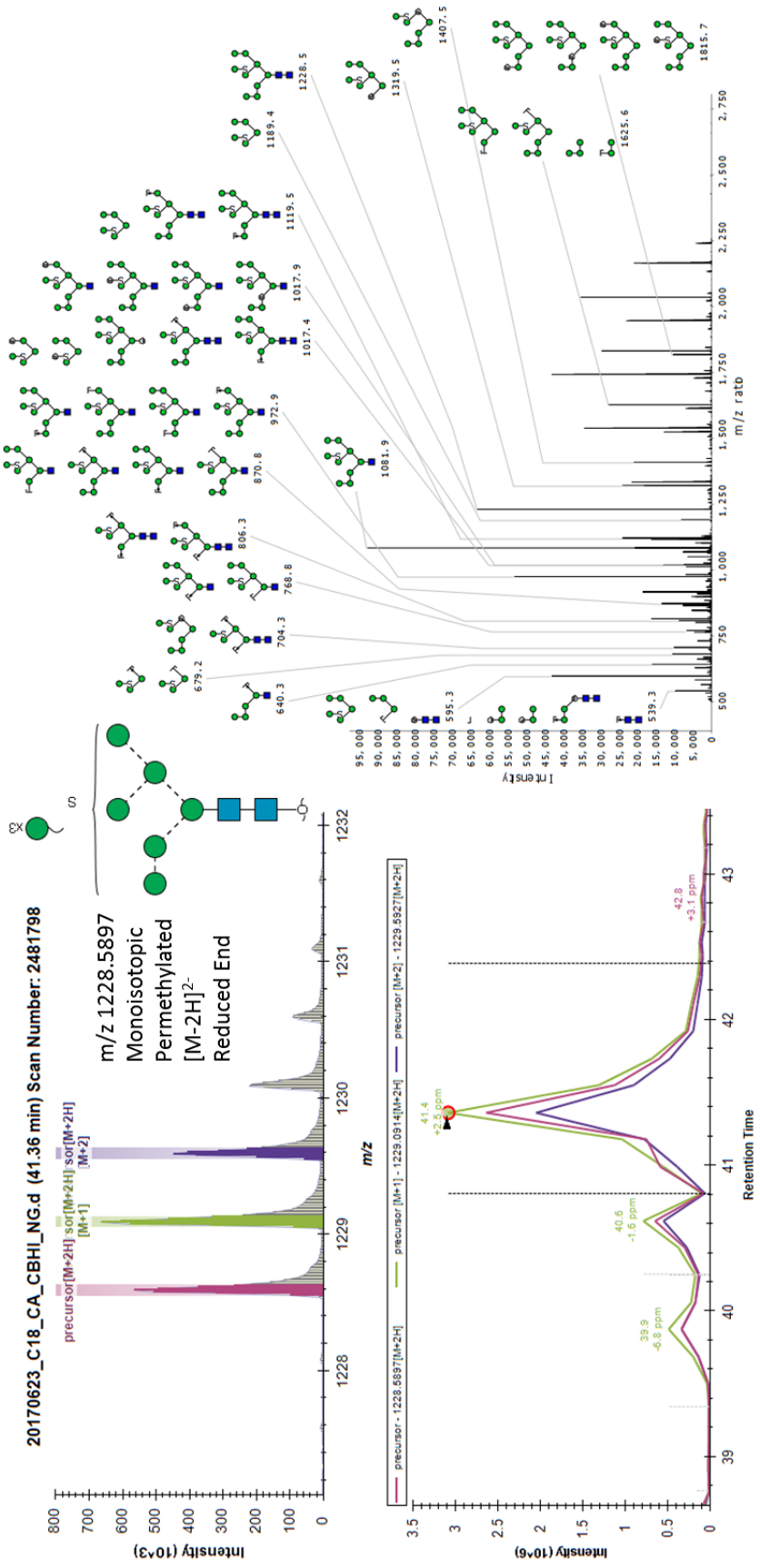
APPENDICES

ETHICS APPROVALS

Ethics approvals (pages 216-230) removed from Open Access version as they may contain sensitive/confidential content.



Supplementary Figure 3.1 Spray stability assessment from direct infusion (5ul/min) of Pierce negative ion calibration solution (monitoring *m/z* 1480.08) using vendor suggested starting tune parameters



Supplementary Figure 4.1 Permethylated glycan analysis and accurate mass analysis to confirm sulfation

		Cell Count Information		Cell Concentration Information(cells/mL)		% of Total Information	
Sample No.	Sample ID_hr	Viable	Dead	Viable	Dead	Viable	Dead
1	u87_lps_57	872	128	3485974.5	511702.66	87.2	12.8
2	hek293_lps_57	879	121	34486816	4747332	87.9	12.1
3	tlr4_lps_57	895	105	27088376	3177966	89.5	10.5
4	u87_lpsrs_57	863	137	3799991.75	603243.19	86.3	13.7
5	hek_lpsrs_57	923	77	18640532	1555060.63	92.3	7.7
6	tlr4_lpsrs_57	875	125	12640535	1805790.75	87.5	12.5
1	u87_lps_68	745	255	2575648.75	881597.88	74.5	25.5
2	hek293_lps_68	897	103	24228508	2782092	89.7	10.3
3	tlr4_lps_68	842	158	27590500	5177314.5	84.2	15.8
4	u87_lpsrs_68	727	273	2413127.75	906167.63	72.7	27.3
5	hek_lpsrs_68	833	167	28711018	5755990.5	83.3	16.7
6	tlr4_lpsrs_68	869	131	23575530	3553963.75	86.9	13.1

Supplementary Table 7.1 Cell viability from serum-free media experiments.

Glycan Class	Glycan Composition (Isomer)	m/z	Adduct	Mean RT	U87MG Relative Glycan Abundance							HEK293 Relative Glycan Abundance						
					0hr	8hr	10hr	12hr	22hr	33hr	44hr	0hr	8hr	10hr	12hr	22hr	33hr	44hr
High Mannose	(Hex)2 + (Man)3(GlcNAc)2	1235.72	[M-H]	33.98	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	1%	1%	1%
	(Hex)3 + (Man)3(GlcNAc)2	698.4	[M-H]	27.04	0%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)4 + (Man)3(GlcNAc)2 (A)	779.44	[M-H]	26.25	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)4 + (Man)3(GlcNAc)2 (B)	779.44	[M-H]	27.27	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)5 + (Man)3(GlcNAc)2 (A)	860.48	[M-H]	26.54	0%	0%	0%	1%	0%	0%	0%	0%	1%	1%	0%	1%	1%	1%
	(Hex)6 + (Man)3(GlcNAc)2	941.52	[M-H]	26.78	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%	0%	3%	2%	2%
Hybrid	(Hex)3 (HexNAc)1 + (Man)3(GlcNAc)2	799.96	[M-H]	33.15	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)1 (NeuAc)1 + (Man)3(GlcNAc)2	864.48	[M-H]	36.35	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%
	(Hex)3 (HexNAc)1 (NeuAc)1 + (Man)3(GlcNAc)2	945.52	[M-H]	36.52	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
Complex	(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2	820.48	[M-H]	36.51	0%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	966.08	[M-H]	33.80	0%	0%	0%	0%	0%	0%	1%	1%	1%	1%	1%	1%	2%	2%
	(Hex)2 (HexNAc)2 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	966.08	[M-H]	41.10	0%	0%	2%	2%	2%	0%	1%	0%	0%	1%	1%	1%	2%	2%
	(Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2 (A)	1111.64	[M-H]	31.77	3%	4%	2%	1%	1%	4%	4%	4%	5%	4%	3%	4%	3%	3%
	(Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2 (B)	1111.64	[M-H]	39.27	7%	6%	3%	2%	3%	7%	7%	7%	6%	7%	6%	7%	6%	6%
	(Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2 (C)	1111.64	[M-H]	45.56	1%	1%	2%	2%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	(Hex)2 (HexNAc)2 (NeuAc)1 (NeuGc)1 + (Man)3(GlcNAc)2 (A)	1119.64	[M-H]	31.51	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (NeuAc)1 (NeuGc)1 + (Man)3(GlcNAc)2 (B)	1119.64	[M-H]	39.14	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%	1%	1%	1%	1%
	(Hex)2 (HexNAc)2 (NeuAc)3 + (Man)3(GlcNAc)2	1257.24	[M-H]	42.33	1%	0%	0%	0%	0%	1%	1%	0%	0%	0%	1%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	893.52	[M-H]	34.22	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (B)	893.52	[M-H]	43.10	0%	0%	7%	7%	10%	1%	1%	0%	0%	0%	0%	1%	1%	1%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	1039.12	[M-H]	41.43	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	1%	2%	2%	2%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	1039.12	[M-H]	47.10	1%	1%	12%	16%	15%	1%	2%	0%	0%	1%	0%	1%	2%	1%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2 (A)	1184.68	[M-H]	41.70	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2 (B)	1184.72	[M-H]	46.90	2%	2%	10%	11%	11%	2%	3%	0%	0%	1%	1%	1%	2%	1%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2 (B)	1184.72	[M-H]	53.03	0%	0%	2%	1%	1%	0%	0%	1%	0%	1%	1%	1%	1%	1%
	(Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2 (A)	1294.24	[M-H]	37.48	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%
	(Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2 (B)	1294.24	[M-H]	42.21	2%	1%	1%	1%	1%	1%	2%	1%	1%	1%	3%	0%	1%	2%
	(Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2 (C)	1294.24	[M-H]	44.73	1%	1%	1%	0%	1%	1%	1%	0%	0%	0%	2%	0%	1%	2%
	(Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2 (D)	1294.24	[M-H]	48.94	0%	1%	1%	0%	1%	0%	1%	0%	0%	0%	2%	0%	1%	1%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (A)	1439.8	[M-H]	39.55	19%	17%	8%	8%	7%	16%	13%	27%	22%	24%	20%	21%	17%	17%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (B)	1439.8	[M-H]	42.47	1%	1%	0%	0%	0%	1%	1%	3%	2%	2%	1%	2%	1%	2%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (C)	1439.8	[M-H]	43.91	2%	2%	1%	1%	1%	2%	2%	0%	0%	1%	2%	0%	0%	0%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (D)	1439.8	[M-H]	46.14	22%	32%	13%	13%	11%	32%	33%	22%	25%	21%	23%	17%	18%	16%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (E)	1439.8	[M-H]	48.46	3%	2%	3%	1%	1%	3%	3%	2%	2%	2%	3%	0%	1%	1%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (F)	1439.8	[M-H]	50.45	1%	1%	1%	0%	0%	0%	1%	0%	0%	0%	0%	0%	1%	0%
	(Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2 (A)	1585.4	[M-H]	36.25	7%	6%	2%	2%	2%	5%	4%	7%	7%	5%	4%	6%	4%	4%
	(Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2 (B)	1585.4	[M-H]	43.10	14%	12%	5%	4%	3%	11%	9%	9%	10%	9%	9%	8%	7%	7%
	(Hex)3 (HexNAc)3 (NeuAc)2 (NeuGc)1 + (Man)3(GlcNAc)2 (A)	1447.8	[M-H]	39.34	2%	2%	1%	1%	1%	2%	1%	3%	2%	2%	2%	2%	2%	2%
	(Hex)3 (HexNAc)3 (NeuAc)2 (NeuGc)1 + (Man)3(GlcNAc)2 (B)	1447.8	[M-H]	45.95	2%	2%	1%	1%	1%	2%	2%	2%	2%	3%	2%	2%	2%	2%
	(Hex)3 (HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	1076.12	[M-H]	43.71	0%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 (HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (B)	1076.12	[M-H]	48.62	0%	0%	0%	1%	1%	0%	0%	0%	1%	0%	1%	0%	0%	0%
	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2	1221.72	[M-H]	45.78	0%	0%	4%	4%	2%	0%	0%	1%	1%	1%	1%	2%	2%	2%

Glycan Class	Glycan Composition (Isomer)	m/z	Adduct	Mean RT	U87MG Relative Glycan Abundance							HEK293 Relative Glycan Abundance						
					0hr	8hr	10hr	12hr	22hr	33hr	44hr	0hr	8hr	10hr	12hr	22hr	33hr	44hr
Complex	(Hex)3 (HexNAc)3 dHex(1) (NeuAc)2 + (Man)3(GlcNAc)2 (A)	1367.32	[M-H]	40.73	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 (HexNAc)3 dHex(1) (NeuAc)2 + (Man)3(GlcNAc)2 (B)	1367.32	[M-H]	45.21	0%	0%	3%	3%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 (HexNAc)3 dHex(1) (NeuAc)2 + (Man)3(GlcNAc)2 (C)	1367.32	[M-H]	48.23	1%	1%	1%	1%	1%	0%	0%	0%	0%	0%	0%	1%	1%	1%
	(Hex)3 (HexNAc)3 dHex(1) (NeuAc)2 + (Man)3(GlcNAc)2 (C)	1367.32	[M-H]	50.30	0%	0%	0%	1%	3%	0%	0%	0%	0%	0%	0%	0%	1%	1%
	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2 (A)	1512.88	[M-H]	46.14	0%	0%	0%	0%	2%	1%	1%	0%	0%	0%	1%	1%	1%	1%
	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2 (B)	1512.88	[M-H]	50.34	1%	1%	3%	4%	1%	0%	0%	1%	1%	1%	1%	1%	1%	1%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 + (Man)3(GlcNAc)2	1258.72	[M-H]	52.35	0%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2	1404.32	[M-H]	51.40	0%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	1%	0%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2 (A)	1549.92	[M-H]	47.82	0%	0%	1%	1%	1%	0%	0%	0%	1%	0%	1%	1%	2%	1%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2 (A)	1695.48	[M-H]	50.50	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2 (B)	1695.48	[M-H]	52.36	0%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2 (C)	1695.48	[M-H]	53.71	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)4 + (Man)3(GlcNAc)2	1227.04	[M-H]	51.38	1%	1%	1%	3%	4%	1%	1%	1%	0%	1%	1%	3%	4%	4%

Supplementary Table 7.2 *N*-glycan structures quantified across time points for U87MG/HEK293 cells with accompanying metadata

Glycan Class	Glycan Composition (Isomer)	m/z	Adduct	Mean RT	PAG Relative glycan abundance											
					No treatment			Sham surgery			Pain treatment (CCI)					
					A8	A10	A13	A1	A2	A6	A11	A4	A7	A9	A12	A14
Hybrid	Hex5 HexNAc3 dHex1 (A)	791.8	[M-H]	33.61	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Hex5 HexNAc3 dHex1 (B)	791.8	[M-H]	40.01	1%	1%	2%	1%	1%	2%	1%	1%	2%	1%	2%	2%
	Hex5 HexNAc3 dHex1 NeuAc1 (A)	937.4	[M-H]	48.41	1%	1%	1%	1%	1%	1%	1%	0%	1%	1%	1%	1%
	Hex5 HexNAc3 dHex1 NeuAc1 (B)	937.4	[M-H]	45.24	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
	Hex6 HexNAc3 dHex1	872.8	[M-H]	38.31	2%	1%	1%	1%	1%	2%	1%	1%	2%	1%	1%	1%
	Hex6 HexNAc3 NeuAc1 (A)	945.3	[M-H]	46.87	1%	0%	0%	0%	0%	1%	1%	0%	1%	0%	1%	1%
	Hex6 HexNAc3 NeuAc1 (B)	945.3	[M-H]	41.73	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Hex6 HexNAc4 dHex1 (A)	974.4	[M-H]	30.89	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	Hex6 HexNAc4 NeuAc1	1047.0	[M-H]	35.91	1%	0%	0%	0%	0%	1%	1%	0%	1%	0%	1%	0%
High Mannose	Hex4 HexNAc2 (A)	1073.5	[M-H]	37.62	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%
	Hex4 HexNAc2 (B)	1073.5	[M-H]	38.67	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Hex5 HexNAc2	1235.4	[M-H]	43.70	10%	13%	11%	13%	11%	11%	27%	11%	10%	18%	12%	11%
	Hex5 HexNAc2	617.2	[M-2H]	43.69	2%	2%	2%	1%	1%	2%	1%	1%	2%	1%	1%	2%
	Hex6 HexNAc2	1397.4	[M-H]	37.85	1%	2%	1%	1%	1%	1%	3%	2%	1%	3%	2%	2%
	Hex6 HexNAc2	698.2	[M-H]	37.87	6%	6%	7%	6%	6%	7%	2%	8%	6%	4%	5%	6%
	Hex7 HexNAc2 (A)	779.3	[M-H]	36.90	6%	6%	6%	6%	6%	6%	5%	6%	7%	6%	7%	6%
	Hex7 HexNAc2 (B)	779.3	[M-H]	38.01	2%	3%	2%	2%	2%	2%	1%	2%	2%	2%	3%	2%
	Hex7 HexNAc2 P1	819.3	[M-H]	30.37	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	1%	1%
	Hex8 HexNAc2	860.3	[M-H]	37.17	9%	8%	7%	9%	9%	7%	7%	9%	8%	11%	7%	8%
	Hex9 HexNAc2	941.4	[M-H]	37.51	9%	8%	7%	12%	12%	9%	12%	13%	9%	12%	7%	9%
	Hex10 HexNAc2	1022.4	[M-H]	39.71	0%	0%	1%	1%	0%	1%	1%	1%	1%	1%	1%	1%
Complex	Hex3 HexNAc3 dHex1	1260.6	[M-H]	50.69	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	0%	0%
	Hex3 HexNAc4 dHex1 (A)	731.3	[M-H]	42.40	12%	11%	13%	9%	12%	11%	2%	8%	12%	6%	10%	10%
	Hex3 HexNAc5	759.8	[M-H]	29.01	3%	3%	3%	2%	2%	4%	0%	2%	4%	1%	4%	4%
	Hex3 HexNAc5 dHex1 (A)	832.8	[M-H]	35.33	14%	13%	11%	19%	16%	15%	15%	17%	12%	15%	14%	14%
	Hex4 HexNAc4 dHex1 (A)	812.3	[M-H]	33.88	2%	2%	2%	2%	1%	2%	1%	2%	2%	1%	3%	2%
	Hex4 HexNAc4 dHex1 (B)	812.3	[M-H]	45.45	1%	1%	3%	1%	1%	1%	1%	1%	2%	1%	1%	2%
	Hex4 HexNAc4 NeuAc1 (A)	884.9	[M-H]	43.12	2%	2%	3%	1%	1%	2%	2%	1%	2%	1%	2%	2%
	Hex4 HexNAc5 dHex1	913.9	[M-H]	29.52	1%	1%	1%	1%	1%	1%	1%	1%	2%	1%	2%	2%
	Hex5 HexNAc4	820.3	[M-H]	30.57	3%	3%	3%	2%	3%	3%	3%	3%	4%	2%	4%	4%
	Hex5 HexNAc4 dHex1	893.3	[M-H]	35.67	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%
	Hex5 HexNAc4 NeuAc1	965.9	[M-H]	33.52	1%	1%	1%	0%	0%	1%	1%	1%	1%	0%	1%	1%
Pauci-mannose	Hex2 HexNAc2	749.3	[M-H]	33.80	1%	1%	1%	1%	1%	0%	1%	1%	0%	1%	1%	1%
	Hex2 HexNAc2 dHex1	895.4	[M-H]	42.23	1%	1%	2%	1%	1%	1%	2%	1%	1%	2%	1%	1%
	Hex3 HexNAc2	911.4	[M-H]	40.73	0%	0%	0%	0%	1%	0%	1%	1%	0%	1%	0%	0%
	Hex3 HexNAc2 dHex1	1057.3	[M-H]	48.27	2%	2%	2%	2%	2%	2%	3%	1%	2%	2%	2%	2%

Glycan Class	Glycan Composition (Isomer)	m/z	Adduct	Mean RT	RVM Relative glycan abundance											
					No treatment				Sham surgery				Pain treatment (CCI)			
					A5	A8	A10	A13	A1	A3	A6	A4	A7	A9	A12	A14
Hybrid	Hex5 HexNAc3 dHex1 (A)	791.8	[M-H]	33.95	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Hex5 HexNAc3 dHex1 (B)	791.8	[M-H]	40.00	1%	1%	1%	1%	1%	1%	1%	2%	1%	1%	2%	2%
	Hex5 HexNAc3 dHex1 NeuAc1 (A)	937.4	[M-H]	48.61	1%	0%	1%	0%	0%	1%	1%	1%	0%	1%	1%	1%
	Hex5 HexNAc3 dHex1 NeuAc1 (B)	937.4	[M-H]	45.66	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
	Hex6 HexNAc3 dHex1	872.8	[M-H]	38.34	1%	1%	1%	1%	2%	2%	1%	2%	2%	1%	2%	2%
	Hex6 HexNAc3 NeuAc1 (A)	945.3	[M-H]	47.09	1%	1%	1%	0%	1%	1%	1%	1%	1%	1%	1%	1%
	Hex6 HexNAc3 NeuAc1 (B)	945.3	[M-H]	39.58	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Hex6 HexNAc4 dHex1 (A)	974.4	[M-H]	33.35	1%	0%	1%	1%	1%	1%	1%	1%	1%	1%	0%	1%
	Hex6 HexNAc4 NeuAc1	1047.0	[M-H]	35.91	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%
High Mannose	Hex4 HexNAc2 (A)	1073.5	[M-H]	37.63	0%	0%	1%	0%	0%	0%	1%	0%	0%	1%	0%	0%
	Hex4 HexNAc2 (B)	1073.5	[M-H]	38.63	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%
	Hex5 HexNAc2	1235.4	[M-H]	43.73	10%	12%	21%	12%	13%	11%	28%	11%	13%	28%	8%	11%
	Hex5 HexNAc2	617.2	[M-2H]	43.69	1%	1%	0%	2%	2%	2%	0%	2%	2%	0%	3%	2%
	Hex6 HexNAc2	1397.4	[M-H]	37.90	1%	2%	3%	1%	1%	1%	4%	1%	1%	4%	1%	2%
	Hex6 HexNAc2	698.2	[M-H]	37.87	7%	7%	3%	9%	7%	7%	2%	7%	7%	2%	9%	7%
	Hex7 HexNAc2 (A)	779.3	[M-H]	36.91	8%	7%	6%	8%	7%	6%	5%	8%	7%	5%	7%	6%
	Hex7 HexNAc2 (B)	779.3	[M-H]	38.05	2%	2%	1%	2%	1%	2%	1%	2%	1%	1%	2%	2%
	Hex7 HexNAc2 P1	819.3	[M-H]	30.40	1%	1%	0%	1%	1%	1%	0%	1%	1%	0%	1%	1%
	Hex8 HexNAc2	860.3	[M-H]	37.20	9%	9%	9%	8%	7%	6%	9%	7%	7%	9%	8%	7%
	Hex9 HexNAc2	941.4	[M-H]	37.57	13%	12%	15%	9%	9%	9%	15%	8%	9%	15%	8%	9%
	Hex10 HexNAc2	1022.4	[M-H]	39.81	0%	1%	1%	1%	0%	1%	1%	1%	0%	1%	1%	1%
Complex	Hex3 HexNAc3 dHex1	1260.6	[M-H]	50.66	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%
	Hex3 HexNAc4 dHex1 (A)	731.3	[M-H]	42.46	5%	4%	3%	9%	8%	8%	2%	9%	8%	2%	8%	8%
	Hex3 HexNAc5	759.8	[M-H]	29.02	2%	2%	1%	2%	3%	4%	1%	3%	3%	1%	4%	4%
	Hex3 HexNAc5 dHex1 (A)	832.8	[M-H]	35.32	19%	21%	15%	15%	15%	14%	9%	14%	15%	9%	17%	14%
	Hex4 HexNAc4 dHex1 (A)	812.3	[M-H]	33.87	2%	2%	1%	2%	3%	3%	1%	3%	3%	1%	2%	3%
	Hex4 HexNAc4 dHex1 (B)	812.3	[M-H]	45.44	1%	0%	0%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	Hex4 HexNAc4 NeuAc1 (A)	884.9	[M-H]	43.17	1%	1%	1%	2%	1%	1%	1%	2%	1%	1%	1%	2%
	Hex4 HexNAc5 dHex1	913.9	[M-H]	29.50	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	Hex5 HexNAc4	820.3	[M-H]	30.59	5%	4%	2%	4%	7%	7%	3%	5%	7%	3%	7%	7%
	Hex5 HexNAc4 dHex1	893.3	[M-H]	35.68	3%	4%	2%	2%	3%	3%	2%	2%	3%	2%	3%	2%
	Hex5 HexNAc4 NeuAc1	965.9	[M-H]	33.49	0%	0%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
Pauci-mannose	Hex2 HexNAc2	749.3	[M-H]	33.84	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	0%
	Hex2 HexNAc2 dHex1	895.4	[M-H]	42.23	1%	1%	2%	1%	1%	1%	2%	1%	1%	2%	1%	1%
	Hex3 HexNAc2	911.4	[M-H]	40.80	0%	0%	1%	1%	0%	0%	1%	0%	0%	1%	0%	0%
	Hex3 HexNAc2 dHex1	1057.3	[M-H]	48.28	1%	1%	2%	1%	1%	1%	3%	1%	1%	3%	1%	2%

Supplementary Table 7.3 *N*-glycan structures quantified from mouse brain sections, PAG and RVM, with accompanying metadata

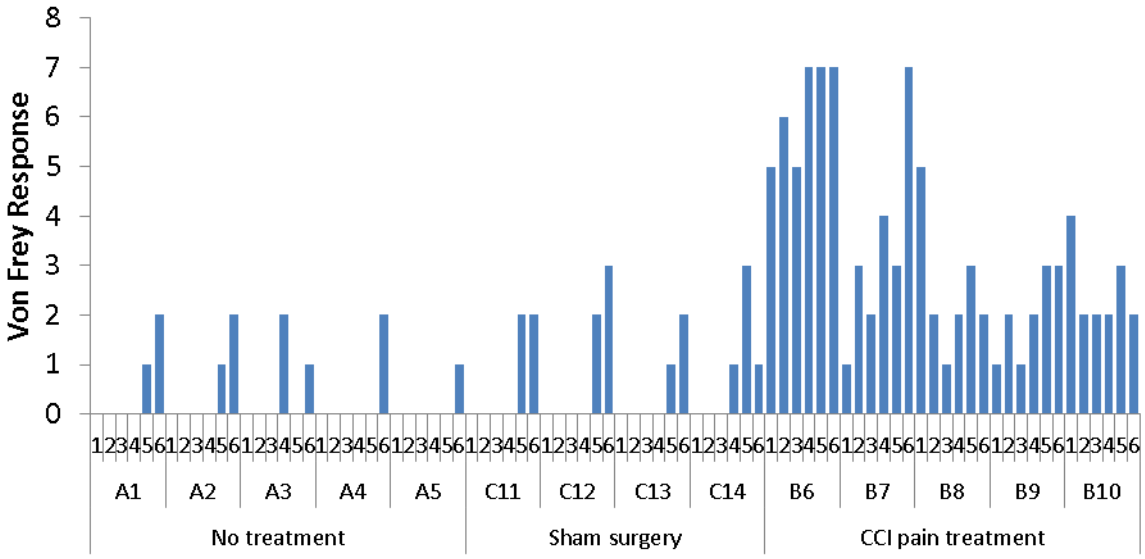
Glycan Class	Glycan Composition (Isomer)	m/z	Adduct	Mean RT	L6 Relative glycan abundance													L5 Relative glycan abundance												L4 Relative glycan abundance								
					No treatment				Sham surgery				Pain treatment (CCI)					No treatment			Sham surgery				Pain treatment (CCI)					No treatment			Sham		Pain treatment (CCI)			
					A5	A8	A10	A13	A1	A2	A3	A6	A4	A7	A9	A12	A14	A5	A10	A13	A1	A2	A3	A6	A4	A7	A9	A12	A14	A8	A10	A13	A3	A6	A4	A7	A9	A12
Hybrid	Hex5 HexNAc3 dHex1 NeuAc1 (A)	937.4	[M-H]	41.05	3%	3%	3%	5%	7%	2%	0%	5%	8%	3%	3%	2%	3%	4%	8%	11%	4%	10%	1%	4%	11%	2%	10%	3%	14%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex5 HexNAc3 dHex1 NeuAc1 (B)	937.4	[M-H]	48.99	2%	1%	2%	2%	7%	5%	4%	3%	4%	2%	2%	2%	1%	9%	4%	5%	2%	6%	11%	3%	6%	2%	5%	1%	6%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex6 HexNAc3 dHex1	872.8	[M-H]	38.28	2%	1%	1%	1%	4%	5%	2%	1%	2%	1%	1%	1%	2%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex6 HexNAc3 NeuAc1 (A)	945.3	[M-H]	36.24	0%	1%	1%	4%	2%	2%	1%	1%	0%	0%	1%	6%	2%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
High Mannose	Hex5 HexNAc2	1235.4	[M-H]	43.88	19%	8%	7%	8%	5%	7%	12%	8%	6%	9%	23%	8%	8%	5%	6%	5%	8%	6%	6%	6%	7%	5%	15%	5%	15%	8%	9%	14%	12%	9%	16%	9%	28%	11%
	Hex6 HexNAc2	1397.4	[M-H]	37.90	3%	1%	1%	1%	0%	1%	2%	1%	2%	1%	4%	0%	2%	1%	1%	1%	2%	1%	1%	1%	1%	2%	3%	1%	3%	2%	2%	2%	3%	2%	3%	2%	4%	2%
	Hex6 HexNAc2	698.2	[M-H]	37.90	2%	8%	6%	6%	4%	6%	6%	7%	4%	9%	1%	6%	7%	3%	8%	5%	5%	4%	4%	8%	5%	9%	2%	7%	1%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex7 HexNAc2 (A)	779.3	[M-H]	36.96	10%	9%	8%	7%	5%	8%	8%	9%	6%	10%	8%	9%	9%	6%	10%	5%	11%	7%	9%	9%	4%	8%	5%	9%	3%	13%	13%	12%	11%	9%	11%	10%	10%	13%
	Hex7 HexNAc2 (B)	779.3	[M-H]	38.00	1%	3%	2%	2%	4%	6%	3%	2%	1%	3%	2%	3%	3%	4%	2%	1%	3%	2%	4%	3%	2%	5%	2%	2%	1%	4%	3%	3%	2%	2%	4%	2%	3%	2%
	Hex8 HexNAc2	860.3	[M-H]	37.28	11%	8%	5%	6%	4%	7%	9%	8%	8%	8%	8%	6%	9%	7%	9%	4%	7%	6%	11%	13%	4%	9%	8%	6%	5%	25%	12%	11%	12%	11%	16%	12%	14%	13%
	Hex9 HexNAc2	941.4	[M-H]	37.62	13%	9%	7%	7%	8%	9%	15%	10%	9%	9%	12%	7%	9%	4%	7%	5%	9%	7%	12%	9%	6%	6%	11%	8%	8%	8%	12%	12%	17%	12%	20%	12%	16%	12%
Complex	Hex3 HexNAc3 dHex1	1260.6	[M-H]	51.45	1%	0%	0%	0%	2%	0%	0%	0%	0%	0%	2%	1%	1%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex3 HexNAc4 dHex1 (A)	731.3	[M-H]	41.68	2%	9%	9%	8%	13%	8%	9%	8%	6%	10%	2%	8%	15%	3%	8%	7%	8%	5%	7%	7%	5%	11%	2%	13%	4%	10%	11%	14%	7%	11%	7%	15%	4%	15%
	Hex3 HexNAc5	759.8	[M-H]	29.37	1%	2%	2%	3%	3%	5%	2%	3%	2%	3%	1%	5%	2%	8%	2%	2%	3%	3%	3%	3%	2%	3%	1%	3%	2%	w	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex3 HexNAc5 dHex1 (A)	832.8	[M-H]	35.46	14%	24%	25%	23%	8%	12%	12%	21%	23%	21%	16%	15%	12%	12%	19%	29%	21%	28%	11%	21%	31%	13%	20%	23%	23%	22%	31%	28%	27%	38%	20%	33%	14%	27%
	Hex4 HexNAc4 dHex1 (A)	812.3	[M-H]	33.96	2%	2%	4%	3%	5%	1%	2%	3%	2%	2%	1%	5%	2%	4%	4%	4%	4%	2%	5%	2%	3%	4%	2%	3%	3%	7%	4%	4%	7%	5%	3%	4%	4%	4%
	Hex4 HexNAc4 dHex1 (B)	812.3	[M-H]	45.37	1%	2%	3%	4%	4%	3%	2%	1%	3%	1%	2%	6%	4%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex4 HexNAc5 dHex1	913.9	[M-H]	37.39	4%	2%	6%	4%	6%	5%	4%	5%	6%	3%	4%	2%	4%	11%	5%	6%	6%	7%	3%	3%	5%	6%	8%	7%	6%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex5 HexNAc4	820.3	[M-H]	30.96	4%	3%	4%	3%	5%	4%	3%	2%	1%	4%	3%	5%	3%	9%	4%	1%	3%	3%	6%	5%	3%	9%	4%	3%	1%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Hex5 HexNAc4 dHex1	893.3	[M-H]	40.14	3%	2%	3%	2%	3%	3%	4%	3%	3%	3%	3%	1%	2%	11%	4%	6%	3%	4%	5%	3%	5%	6%	4%	4%	4%	2%	3%	2%	3%	1%	1%	3%	3%	1%
Pauci-mannose	Hex3 HexNAc2 dHex1	1057.3	[M-H]	48.23	3%	1%	1%	1%	1%	2%	2%	1%	1%	1%	2%	1%	2%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Supplementary Table 7.4 First round of quantification of Spinal cord *N*-glycans with accompanying metadata

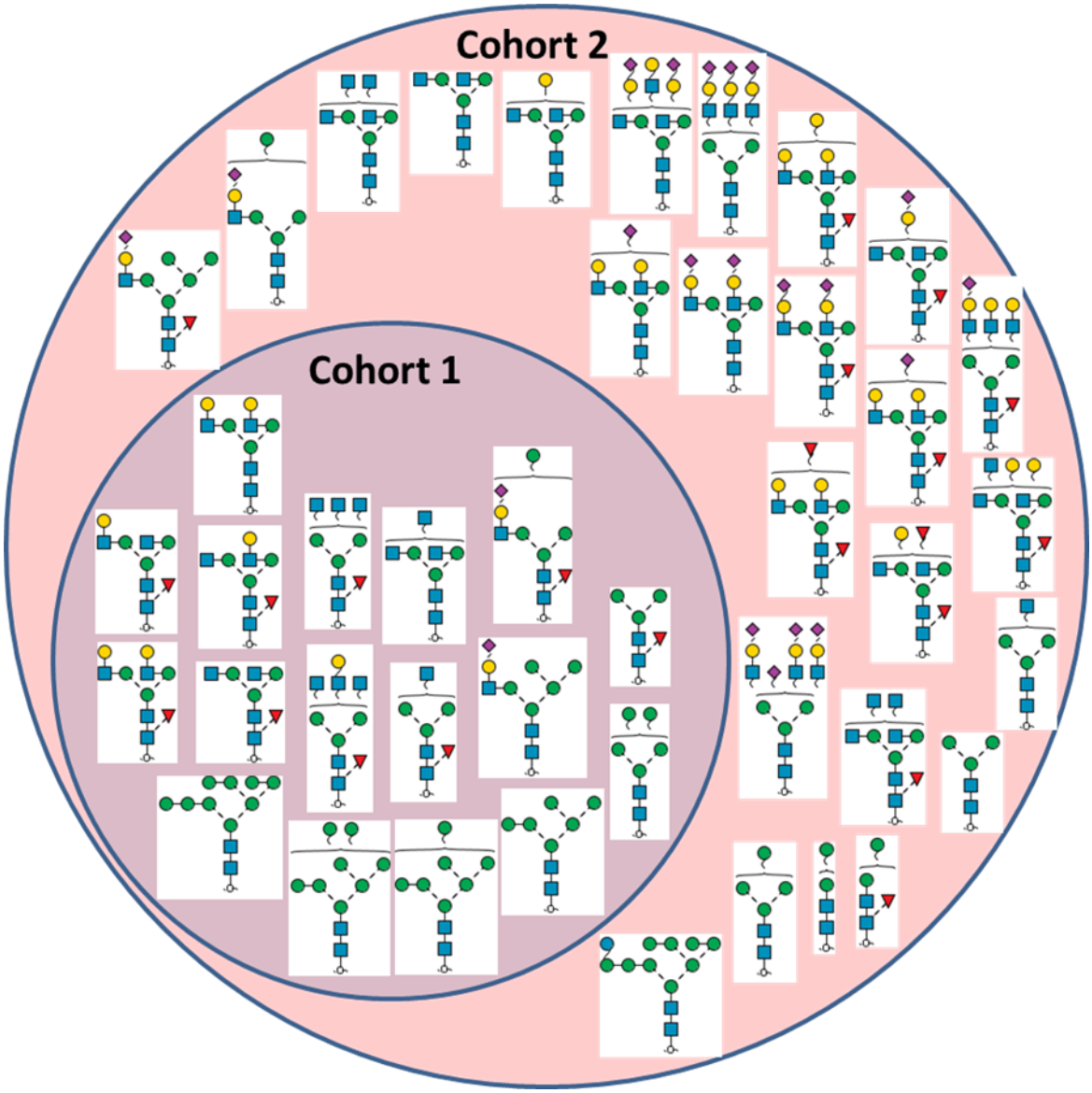
Glycan Class	Glycan Composition	m/z	Adduct	Mean RT	LHS Relative glycan abundance									RHS Relative glycan abundance								
					Control			Sham surgery			CCI			Control			Sham			CCI		
					A20	A21	A22	A1	A2	A5	A3	A4	A11	A20	A21	A22	A1	A2	A5	A3	A4	A11
High mannose	(Hex)1 + (Man)3(GlcNAc)2 (A)	1236	[M-H]	47.15	9%	11%	10%	13%	10%	9%	12%	11%	12%	10%	10%	9%	9%	11%	9%	10%	9%	11%
	(Hex)2 + (Man)3(GlcNAc)2 (A)	1073	[M-H]	41.27	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 + (Man)3(GlcNAc)2 (A)	698.2	[M-2H]	40.93	5%	4%	5%	6%	5%	5%	4%	4%	4%	5%	4%	6%	5%	4%	5%	5%	4%	5%
	(Hex)4 + (Man)3(GlcNAc)2 (A)	779.3	[M-2H]	39.31	6%	6%	6%	4%	5%	6%	5%	6%	4%	6%	6%	5%	7%	4%	6%	6%	4%	7%
	(Hex)4 + (Man)3(GlcNAc)2 (B)	779.3	[M-2H]	40.93	2%	2%	1%	1%	2%	2%	2%	2%	1%	2%	1%	2%	2%	2%	2%	2%	2%	3%
	(Hex)5 + (Man)3(GlcNAc)2 (A)	860.3	[M-2H]	39.66	5%	8%	7%	6%	6%	6%	4%	7%	8%	7%	7%	6%	7%	6%	7%	6%	6%	6%
	(Hex)6 + (Man)3(GlcNAc)2 (A)	941.4	[M-2H]	40.01	11%	13%	10%	12%	11%	14%	11%	9%	10%	10%	10%	11%	11%	9%	11%	11%	12%	8%
	(Hex)7 + (Man)3(GlcNAc)2 (A)	1022	[M-2H]	42.31	1%	1%	1%	1%	1%	0%	1%	1%	1%	0%	1%	1%	0%	1%	1%	1%	1%	1%
Hybrid	(Hex)2 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	937.4	[M-2H]	43.58	1%	1%	1%	1%	3%	1%	1%	2%	0%	1%	3%	2%	1%	3%	1%	1%	2%	3%
	(Hex)2 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	937.4	[M-2H]	51.06	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	2%	1%	1%	1%	1%
	(Hex)2 (HexNAc)1 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	791.8	[M-2H]	37.24	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)1 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (B)	791.8	[M-2H]	43.46	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	(Hex)3 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	1018	[M-2H]	46.46	0%	1%	1%	0%	1%	1%	0%	1%	1%	1%	1%	1%	1%	1%	1%	0%	1%	0%
	(Hex)3 (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	1018	[M-2H]	49.8	1%	1%	1%	1%	1%	1%	0%	1%	1%	1%	1%	0%	1%	0%	1%	1%	1%	1%
	(Hex)3 (HexNAc)1 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	872.8	[M-2H]	41.5	1%	1%	2%	1%	1%	1%	1%	1%	1%	2%	1%	1%	1%	1%	1%	2%	1%	2%
	(Hex)3 (HexNAc)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	945.3	[M-2H]	42.19	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%
	(Hex)3 (HexNAc)1 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	945.3	[M-2H]	49.34	0%	1%	0%	1%	1%	1%	1%	0%	1%	1%	1%	0%	1%	1%	1%	1%	1%	1%
Complex Bi	(Hex)1 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	957.8	[M-2H]	46.57	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	(Hex)1 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	957.8	[M-2H]	50.37	1%	1%	1%	0%	0%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%	1%	1%	1%
	(Hex)1 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	957.8	[M-2H]	54.64	1%	1%	1%	0%	0%	1%	1%	0%	1%	1%	1%	1%	1%	1%	0%	1%	1%	1%
	(Hex)1 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	812.3	[M-2H]	37.01	1%	2%	2%	2%	2%	2%	2%	2%	1%	2%	1%	2%	2%	2%	1%	2%	1%	2%
	(Hex)1 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (B)	812.3	[M-2H]	48.65	1%	1%	1%	1%	1%	1%	1%	0%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	(Hex)1 (HexNAc)2 (Deoxyhexose)2 + (Man)3(GlcNAc)2 (A)	885.3	[M-2H]	46.46	3%	2%	2%	2%	1%	2%	2%	2%	3%	2%	1%	3%	1%	1%	2%	3%	2%	2%
	(Hex)1 (HexNAc)2 + (Man)3(GlcNAc)2 (A)	739.3	[M-2H]	31.13	1%	1%	1%	1%	0%	1%	1%	0%	0%	1%	1%	1%	1%	0%	0%	0%	0%	1%
	(Hex)1 (HexNAc)2 + (Man)3(GlcNAc)2 (B)	739.3	[M-2H]	33.44	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	1039	[M-2H]	44.84	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	1039	[M-2H]	47.84	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%

Glycan Class	Glycan Composition	m/z	Adduct	Mean RT	LHS Relative glycan abundance											RHS Relative glycan abundance										
					Control			Sham surgery			CCI					Control			Sham			CCI				
					A20	A21	A22	A1	A2	A5	A3	A4	A11	A20	A21	A22	A1	A2	A5	A3	A4	A11
Complex Bi	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 +	1184	[M-2H]	46.11	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2 (B)	1184	[M-2H]	49.11	1%	0%	0%	1%	0%	1%	1%	1%	0%	1%	1%	0%	0%	0%	0%	1%	1%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2 (C)	1184	[M-2H]	52.45	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	893.3	[M-2H]	31.13	2%	2%	2%	2%	3%	2%	2%	2%	2%	2%	3%	1%	2%	2%	2%	3%	3%	2%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	893.3	[M-2H]	38.97	2%	2%	2%	2%	2%	2%	2%	3%	2%	3%	2%	2%	2%	2%	2%	2%	2%	2%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	893.3	[M-2H]	41.96	0%	1%	1%	0%	1%	1%	1%	0%	0%	1%	1%	0%	1%	1%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)2 + (Man)3(GlcNAc)2 (A)	966.3	[M-2H]	36.66	2%	2%	2%	2%	1%	2%	2%	1%	2%	2%	2%	1%	2%	1%	2%	2%	2%	2%
	(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2 (A)	820.3	[M-2H]	33.55	2%	3%	3%	2%	3%	4%	3%	2%	3%	3%	2%	2%	2%	4%	3%	3%	2%	3%
	(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2 (B)	820.3	[M-2H]	37.01	1%	0%	1%	1%	1%	1%	1%	0%	1%	1%	1%	1%	0%	0%	0%	1%	1%	1%
	(Hex)3 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	974.4	[M-2H]	33.9	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
	(Hex)3 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (B)	974.4	[M-2H]	44.61	0%	1%	0%	1%	0%	0%	1%	0%	1%	0%	0%	1%	1%	1%	1%	1%	1%	1%
	HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	731.2	[M-2H]	46	8%	5%	6%	5%	6%	6%	7%	4%	7%	5%	4%	6%	7%	6%	5%	5%	7%	5%
Complex Tri	(Hex)1 (HexNAc)3 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (A)	1059	[M-2H]	43.46	2%	1%	1%	1%	1%	1%	1%	1%	2%	1%	1%	2%	1%	1%	1%	1%	1%	1%
	(Hex)1 (HexNAc)3 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2 (B)	1059	[M-2H]	45.65	1%	1%	1%	1%	2%	1%	1%	1%	1%	1%	1%	1%	1%	2%	1%	0%	2%	0%
	(Hex)1 (HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	913.8	[M-2H]	40.93	2%	2%	2%	2%	2%	3%	2%	2%	1%	2%	2%	3%	2%	3%	2%	2%	2%	2%
	(Hex)2 (HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	994.8	[M-2H]	42.42	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2 (A)	959.4	[M-2H]	54.64	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 (HexNAc)3 + (Man)3(GlcNAc)2 (A)	759.8	[M-2H]	32.29	1%	2%	2%	1%	1%	2%	2%	1%	2%	2%	2%	2%	2%	2%	2%	1%	1%	1%
	(HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	832.8	[M-2H]	39.08	12%	10%	13%	12%	11%	10%	14%	17%	14%	11%	15%	15%	13%	10%	15%	15%	13%	12%
Complex Tetra	(HexNAc)4 (Deoxyhexose)1 + (Man)3(GlcNAc)2 (A)	934.5	[M-2H]	39.2	1%	2%	1%	1%	1%	1%	1%	2%	1%	1%	1%	1%	1%	1%	1%	2%	1%	1%
	(HexNAc)4 + (Man)3(GlcNAc)2 (A)	861.3	[M-2H]	39.66	2%	3%	2%	2%	3%	2%	2%	3%	3%	3%	2%	3%	3%	3%	2%	3%	2%	2%
Pauciman nose	(Hex)1 (HexNAc)2 (A)	587.3	[M-H]	35.86	1%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (A)	749.2	[M-H]	37.59	0%	0%	0%	1%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)2 (HexNAc)2 (Deoxyhexose)1 (A)	895.3	[M-H]	46.11	2%	1%	0%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%	0%
	(Hex)3 (HexNAc)2 (A)	911.4	[M-H]	44.15	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	(Hex)3 (HexNAc)2 (Deoxyhexose)1 (A)	1057	[M-H]	51.41	2%	1%	1%	1%	1%	1%	1%	1%	2%	1%	1%	1%	2%	1%	2%	1%	1%	1%

Supplementary Table 7.5 Characterised structures corresponding to quantified *N*-glycans from additional spinal cord material



Supplementary Figure 7.1 Behavioural Von-Frey data for each animal



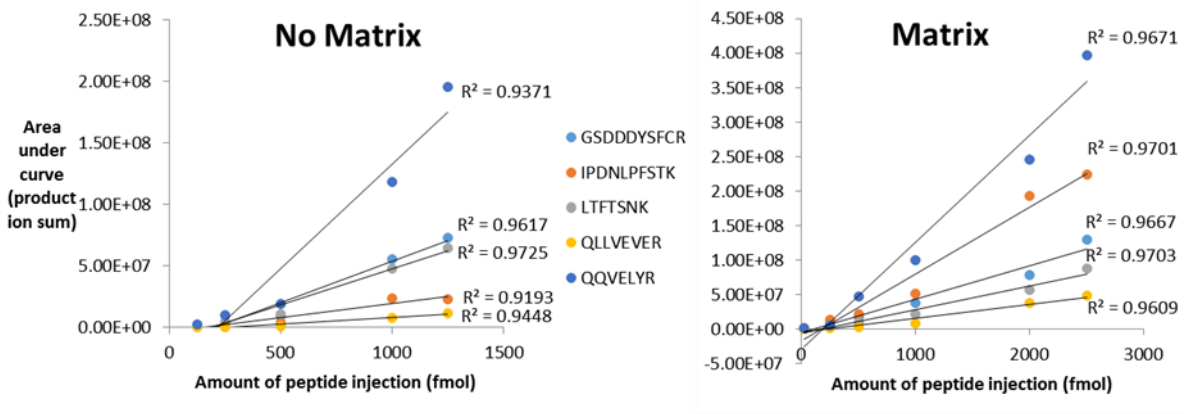
Supplementary Figure 7.2 Expansion of spinal cord detected N-glycans using more tissue and structural comparison to initial low sample *N*-glycan analysis

Band	MD-2 peptides detected (0.05 FDR)	Coverage
2 (~20 kDa)	29	67.5%
3 (~25 kDa)	32	58.1%
4 (~30 kDa)	6	26.3%

Supplementary table 8.1 Bottom-up proteomics results from analysis of SDS-PAGE bands 2,3,4

Peptide spectral matches	Protein Name
61	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2
60	Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE=1 SV=1
58	Tubulin beta-4A chain OS=Homo sapiens GN=TUBB4A PE=1 SV=2
37	Tubulin beta-2B chain OS=Homo sapiens GN=TUBB2B PE=1 SV=1
22	78 kDa glucose-regulated protein OS=Homo sapiens GN=HSPA5 PE=1 SV=2
18	Protein phosphatase 1B OS=Homo sapiens GN=PPM1B PE=1 SV=1
15	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=1
12	Endoplasmin OS=Homo sapiens GN=HSP90B1 PE=1 SV=1
12	Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4
11	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1
11	Dynein heavy chain 9, axonemal OS=Homo sapiens GN=DNAH9 PE=1 SV=3
9	60 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPD1 PE=1 SV=2
8	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1
8	Histone H1.3 OS=Homo sapiens GN=HIST1H1D PE=1 SV=2
8	Endoplasmic reticulum resident protein 44 OS=Homo sapiens GN=ERP44 PE=1 SV=1
8	Sodium/potassium-transporting ATPase subunit alpha-1 OS=Homo sapiens GN=ATP1A1 PE=1 SV=1
8	DNA-dependent protein kinase catalytic subunit OS=Homo sapiens GN=PRKDC PE=1 SV=3
8	Protein furry homolog OS=Homo sapiens GN=FRY PE=1 SV=1
7	Elongation factor Tu, mitochondrial OS=Homo sapiens GN=TUFM PE=1 SV=2
7	Lymphocyte antigen 96 OS=Homo sapiens GN=LY96 PE=1 SV=2

Supplementary table 8.2 Peptides detected from in-solution trypsin digestion of recombinant MD-2 (aka LY96)



Supplementary figure 8.1 Standard curve based on total ion abundance for b and y product ions, instead of precursor monoisotopic peak for each respective peptide