Preparative Scale Release and Purification of *N*-linked Glycan Standards from Human Milk Glycoproteins

Raymond Wei Wern Chong

B. Biotech., B. Sc. (Hons)

This thesis is presented for the Master of Philosophy in Chemistry and Biomolecular Sciences

Department of Chemistry and Biomolecular Sciences

Faculty of Science

Macquarie University

Sydney, NSW

February 2014

Table of Contents

Abstract

Glycosylation has been implicated in many important biological processes, disease states, and host-pathogen interactions yet progress in understanding the biological roles and functions of glycosylation has been limited by the lack of standardised tools and techniques. Purified glycans of specific structure are extremely valuable as standards for structural characterisation of glycans as well as for the calibration of instruments, and the validation or comparison of analytical techniques. However, structurally defined glycans have proven difficult to obtain in ample quantities, are extremely complex to produce synthetically, and when commercially available, are considerably expensive.

In this work, a method was developed for obtaining preparative scale quantities of specific *N*-linked glycan structures of natural origin in a laboratory based setting, as a means to prepare glycan standards for analysis and other downstream processes. Human milk was selected as a biologically relevant source of *N*-linked glycans due to the high glycoprotein content and growing interest in the bioactivity of human milk glycoconjugates. A protocol for the preparative scale release of *N*-linked glycans from human milk proteins and a separation technique using high performance liquid chromatography (HPLC) on porous graphitised carbon (PGC) was developed. A method of offline detection was also established using matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) to identify *N*-linked glycans in the fractions separated by PGC-HPLC.

Six structurally unique *N*-linked glycans from human milk were separated and collected at high purity (~90%). Characterisation by PGC-LC coupled to electrospray ionisation tandem mass spectrometry (PGC-LC-ESI-MS/MS) identified a set of fucosylated *N*-linked glycans, structures which are not currently commercially available. Fucosylation is a specific feature of human glycosylation that is of recent interest to the medical research field, hence the

development of a preparative scale technique to obtain fucosylated glycans is of significant value. A retention time library and MS fragmentation dataset of fucosylated glycans was established to facilitate their use as standards. In addition, the application of the human milk *N*-linked glycans in an *in vitro* assay was investigated, in which the glycans were assessed for their potential to inhibit bacterial binding to membrane proteins extracted from human intestinal cells.

Declaration

The research presented in this thesis is an original study and was completed by the author, a postgraduate student completing the Masters of Philosophy, under the supervision of Professor Nicolle H. Packer. This study was completed in the Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW, Australia.

I certify that any help received in preparing this thesis and all sources used have been acknowledged. Furthermore, I also certify that the work presented in this thesis has not been submitted to any other university or institution for any degree or qualification.

Raymond Wei Wern Chong

Acknowledgements

First and foremost I would like to thank my supervisor, Professor Nicolle H. Packer, for all of her cheerful support and expert advice, the gentle nudges in the right direction, and being able to make me feel significantly better leaving her office than when entering.

I would especially like to thank Dr. Robyn Peterson who provided extensive linguistic support by thoroughly proofreading the many drafts of this thesis.

Thank you to Dr. Liisa Kautto for her laboratory expertise and the postdocs past and present, Morten, Maja, Arun, Terry and Chi-hung for their knowledge and guidance.

I would also like to thank the members of the GlycoMQ group for providing a sense of community, a source of creativity, and much needed feedback. Special thanks go out to Jodie, who was so supportive and always willing to lend a helping hand.

Finally, I would like to thank my family who have supported me unconditionally throughout this project.

.

Abbreviations

ACN	Acetonitrile
Asn	Asparagine
BPC	Base Peak Chromatogram
BSA	Bovine Serum Albumim
CE	Capillary Electrophoresis
CFG	Consortium for Functional Glycomics
CID	Collision Induced Dissociation
CNX	Calnexin
CRT	Calreticulin
DHB	Di-hydroxybenoic acid
DMSO	Di-methyl sulfoxide
DTT	Di-thiolthreitol
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
ER	Endoplasmic reticulum
ESI	Electrospray Ionisation
ESI-MS	Electrospray ionization mass spectrometry
FBS	Fetal bovine serum
GBP	Glycan binding protein
GDP	Guanosine diphosphate
GPC	Gel Permeation Chromatography
GT	Glycosyltransferase
HCl	Hydrochloric Acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HILIC	Hydrphilic Interaction Liquid Chromatography
HPAEC-PAD	High Performance Anion Exchange Chromatography Pulsed Amperometric Detection
HPLC	High performance liquid chromatographic
i.d.	Interal Diameter
IAA	Iodoacetamide
LB	Luria broth
LC	Liquid Chromatography

LC-MS	Liquid Chromatography-Mass Spectrometry
Le^a	Lewis a epitope
Le^b	Lewis b epitope
Le^x	Lewis x epitope
Le^y	Lewis y epitope
MALDI-TOF	Matrix assisted laser desorption/ionization – time-of-flight
MeOH	Methanol
MFGM	Milk fat globule membrane
MS	Mass Spectrometry, Mass Spectrometer
NaOH	Sodium Hydroxide
PBS	Phosphate buffered saline
PGC	Porous graphitized carbon
PNGase F	Peptide N glycosidase F
ppGalNAcT	Polypeptide-N-acetylgalactosaminetransferases
PVDF	Polyvinylidene fluoride
RP	Reversed Phase
SDS	Sodium dodecyl sulphate
Ser	Serine
SEC	Size Exclusion Chromatography
SPE	Solid phase extraction
TFA	Trifluoroacetic acid
THAP	2',4',6'-trihydroxyacetophenone
Thr	Threonine
TSB	Tryptone soya bean
UDP	uridine diphosphate
UV	Ultraviolet

Monosaccharides

Fuc	Fucose
Gal	Galactose
GalNAc	*N*-acetylgalactosamine
Glc	Glucose
GlcA	Glucuronic acid
GlcNAc	*N*-acetylgalactosamine
IdoA	Iduronic acid
Man	Mannose
NeuAc	*N*-acetylneuraminic acid
NeuGc	*N*-glycolylneuraminic acid
Xyl	Xylose

Chapter 1 Introduction

Glycosylation describes the attachment of oligosaccharides or "glycans" to biological molecules such as proteins and lipids, a process now known to be ubiquitous in nature across all domains of life from eukaryotes, to bacteria and the Archaea (Lauc and Zoldos, 2010). Glycosylation has been implicated in a number of important biological processes such as regulation and modification of protein folding, function and resistance to degradation (Dwek, 1996), cellular communication (Boscher et al., 2011) and host-pathogen interactions (Varki, 1993). Significant variations in glycosylation have also been associated with a number of common disease states including cancer (Hakomori, 1996; Reis et al., 2010), diabetes (Brownlee, 1995), and cystic fibrosis (Scanlin and Glick, 1999). However, progress in understanding the functions of glycosylation has been limited by the remarkable complexity and heterogeneity exhibited by glycans (Lauc and Zoldos, 2010).

The drive to investigate glycosylation has resulted in the establishment of glycobiology and glycomics as emerging fields of research (Rademacher et al., 1988; Varki, 2009). The term glycobiology was first coined by Raymond Dwek in 1988 and involves the study of glycans in order to understand their structure, chemistry, biosynthesis and biological functions (Rademacher et al., 1988; Varki, 2009). As a subset of glycobiology, glycomics involves the systematic analysis of glycan structures that occur as free oligosaccharides or conjugated to proteins and lipids as the "glycome" of cells, tissues or organisms (Varki, 2009). Since glycomics is still a developing field of research, the work presented in this thesis aims to contribute to the development of glycomic analysis of glycoproteins in particular, by the preparation of glycan standards of known structure from a natural human source, breast milk. This chapter provides an introduction to the context of this work. The principles of glycobiology are outlined and methods currently used in glycomics are discussed. The glycosylated structures in human milk are described in greater detail in Chapter 3.

1.1 Principles of Glycobiology

The widespread occurrence of glycosylation on cytosolic, secreted and cell membrane proteins is suggestive of the involvement of glycan epitopes in a vast range of biological processes (Varki, 2009). The cells of higher eukaryotes are covered with a dense layer of glycoconjugates (glycocalyx) that displays important cellular information and provides a potential target for pathogenesis (Lauc and Zoldos, 2010). Glycoconjugates with glycan epitopes similar to those in the glycocalyx are also present in secreted fluids such as milk, saliva and tears, which appear to play a role in disease prevention (Peterson et al., 2013). The complexity of glycans, biosynthesis and conjugation to proteins is described below.

1.1.1 Glycans are Complex and Diverse in Structure

Oligosaccharides or glycans are defined as polymers comprised of monosaccharide monomers (Varki, 2009). Monosaccharides are polyhydroxy aldehydes or polyhydroxy ketones containing five (pentose), six (hexose) or nine (neuraminic acid) carbon atoms (Marino et al., 2010). The structural diversity of monosaccharides arises from the high number of stereogenic centres resulting in myriad of potential stereoisomers. For example, many monosaccharides are epimers (differ in the configuration of a single carbon atom); glucose, galactose, and mannose are epimeric forms of hexose. Furthermore, monosaccharides can exist as the D- or L-enantiomers (mirror images that are non-superimposable), which are recognised as biologically distinct from one another resulting in unique specificities and activities (Figure 1.1)(Varki, 2009). Within human systems, only ten monosaccharides are commonly found in glycoconjugates. These are glucose (Glu), galactose (Gal), mannose (Man), xylose (Xyl), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), fucose (Fuc), glucuronic acid (GlcA), iduronic acid (IdoA), *N*-acetylneuraminic acid (NeuAc). Vertebrates generally have D-monosaccharides with the exception of L-fucose and L-iduronic acid (Varki, 2009).
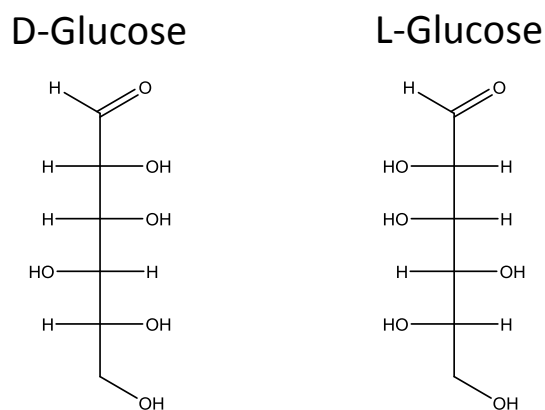
Figure 1.1 – Comparison of the open chain D- and L-enantiomers of glucose represented as Fischer projections.

Open chain monosaccharides can undergo intramolecular rearrangement to form cyclic hemiacetals resulting in six (pyranosidic) or five (furanosidic) membered rings (Figure 1.2) (Marino et al., 2010). When in cyclic form, two anomers become possible (α and β) due to the formation of a new chiral centre known as the anomeric carbon (Figure 1.2, indicated by arrows). Anomericity is determined by the stereochemical configuration of the anomeric (C1) carbon and the highest numbered chiral centre (Figure 1.1.)(Marino et al., 2010). If both carbons are fixed in opposite directions then anomericity is α; if both carbons are fixed in the same direction then anomericity is β.
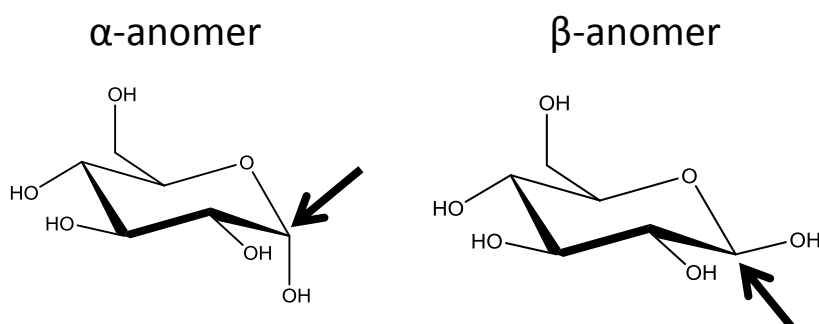
Figure 1.2 – The configuration of α and β anomers of glucose in the pyranosidic (6 membered ring) form represented in the chair conformation. The anomeric carbon (C1) is indicated by arrows.

Monosaccharides are bonded together by reaction of the anomeric carbon with any hydroxyl group of another monosaccharide in the formation of a glycosidic bond (Marino et al., 2010). Since a number of hydroxyl groups are available, many linkage positions are possible and

further influenced by the configuration of the anomeric carbon; just two hexoses can be linked via C1 to C2, C3, C4 or C6 with either α or β linkage resulting in eight possible configurations. Due to the acidic nature of neuraminic acid (NeuAc), glycans containing NeuAc are considered acidic glycans while glycans without NeuAc are considered neutral glycans. Additional modifications such as sulfation, acetylation and phosphorylation result in an incredibly diverse range of glycan structures each with a unique spectrum of bioactivity (Varki, 2009).

1.1.2 Symbol Notation Systems to Represent Glycans

Due to the structural complexity of monosaccharides, two main systems have been developed to simplify the visual representation of glycans. The system developed by the consortium of functional genomics (CFG) represents monosaccharides as coloured shapes (Figure 1.3).
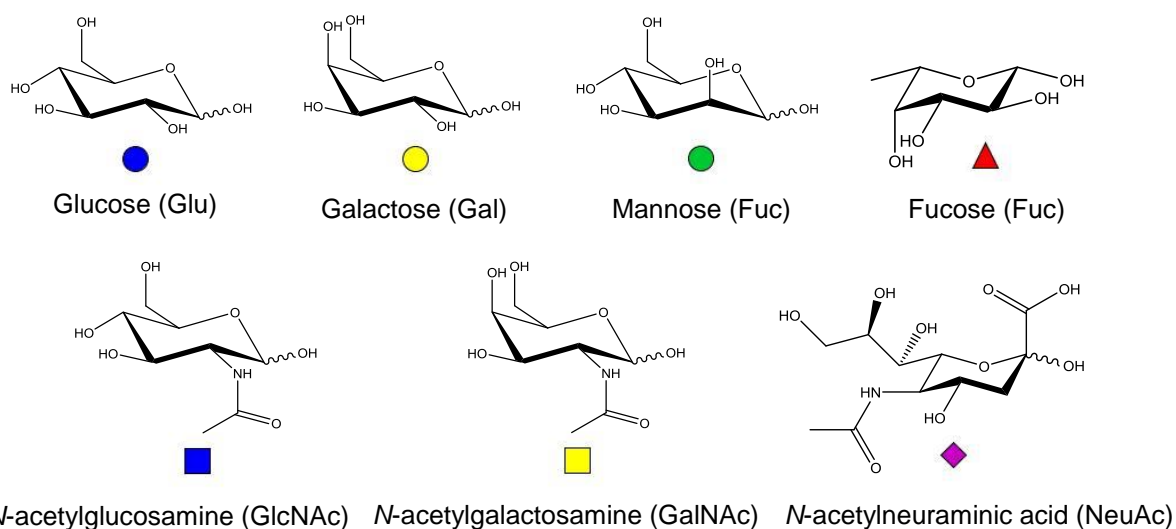
Figure 1.3 – The most commonly found monosaccharides in humans represented as full chemical structures accompanied by the corresponding CFG symbol (Ceroni et al., 2008). The wavy lines indicate both α and β anomers. Fucose is shown as the L-enantiomer.

Hexoses are depicted as circles and *N*-acetylhexoses are depicted as squares with epimers distinguished by colour; fucose and *N*-acetylneuraminic acid are represented by a red triangle and purple diamond respectively. The Oxford symbol notation also represents monosaccharides as symbols but gives additional information on linkage types and positions by the placement of the glycan symbols in the figure (Figure 1.4) (Harvey et al., 2009).

16
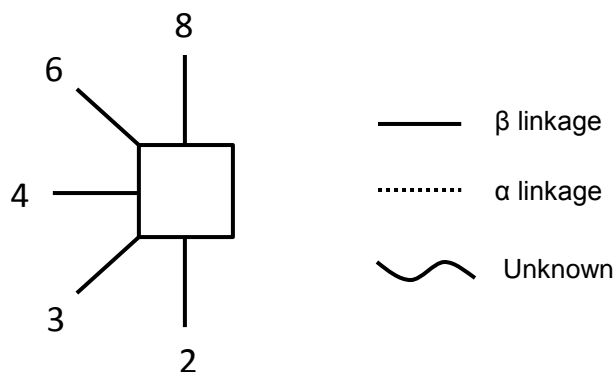
Figure 1.4 –The Oxford linkage notation indicating the position of monosaccharide substitution and type of linkage (Harvey et al., 2009).

A hybrid notation system was used in this thesis for the representation of glycan structures. The coloured symbols of the CFG format were combined with the Oxford system in which residues are placed with respect to linkage position when known, while α or β linkage is defined on the monosaccharide containing the linking anomeric carbon.

1.1.3 Protein Glycosylation

Glycosylation of proteins is the most common post-translational modification and it has been estimated that over 70% of proteins are glycosylated (Apweiler et al., 1999). Two main types of protein glycosylation are found in nature, distinguished and named by the mechanism of attachment to the amino acids of a protein (Varki, 2009). Glycosylation of the hydroxyl group (via the oxygen) of serine (Ser) or threonine (Thr) residues results in the attachment of *O*-linked glycans. There are eight possible *O*-linked core structures of which core 1 to 4 are most commonly found in humans (Figure 1.5) (Varki, 2009). *O*-linked glycosylation is most prevalent in mucins, a class of high molecular weight glycoproteins that contain a variable number of tandem repeat (VNTR) regions rich in proline, serine and threonine, of which the latter two are highly glycosylated (Van Klinken et al., 1998).
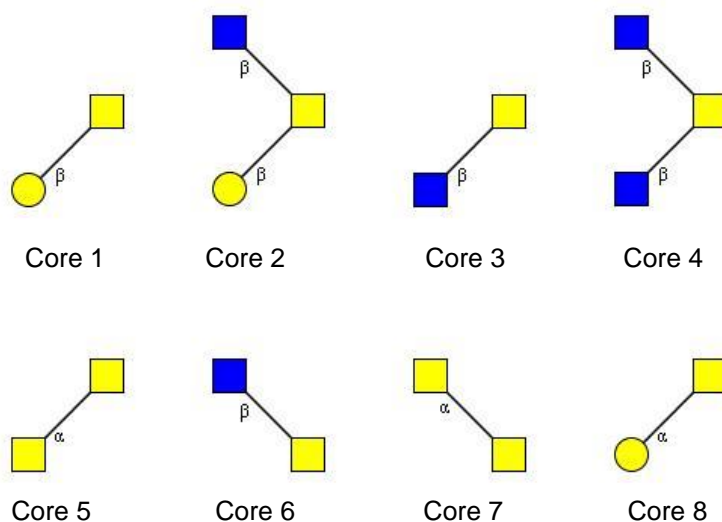
Figure 1.5 – The eight core structures of *O*-linked glycans, core 1 – 4 have widespread occurrence while core 5 and 8 have restricted occurrence and core 6 and 7 are not found in humans (Varki, 2009).

Glycosylation of the amide group (via the nitrogen) of an asparagine (Asn) residue results in the attachment of *N*-linked glycans. There are three classes of *N*-linked glycans which share a conserved core structure known as the chitobiose core (Figure 1.6)(Varki, 2009). *N*-linked glycosylation generally requires an Asn-X-Ser/Thr consensus sequence where X may be any amino acid except proline, thus allowing the prediction of potential *N*-linked glycosylation sites, which is not possible for *O*-linked glycosylation sites.

N-linked glycosylation is more prevalent than *O*-linked glycosylation; an analysis of known glycoproteins in the SWISS-PROT database found that 16% were only *O*-glycosylated while the remaining 84% contained both *N*-linked as well as *O*-linked glycosylation with an average of three potential *N*-linked glycosylation sites per protein (Apweiler et al., 1999). Taken together with the fact that that there is an enzyme that removes all mammalian *N*-linked glycans and many glycobiology research groups focus on *N*-linked glycosylation and require *N*-linked glycan standards for analysis, this work presented in this thesis is specifically focused on *N*-linked glycans.
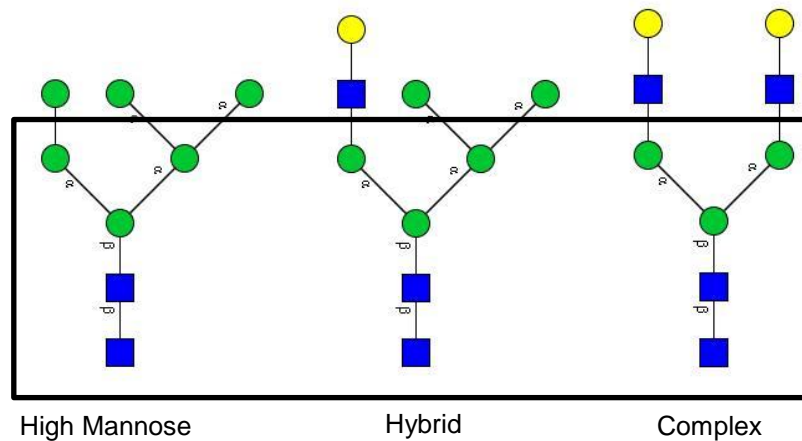
Figure 1.6 – The three classes of *N*-linked glycans, the common chitobiose core is boxed (Varki, 2009).

Protein glycosylation results in a number of possible *glycoforms* of a single glycoprotein; each *glycoform* contains a different set of glycans attached to each glycosylation site on the protein (Rademacher et al., 1988). The diversity of possible *glycoforms* thus results in a glycoproteome that is several orders of magnitude more complex than the proteome itself, which significantly expands the flexibility and potential functions of an otherwise finite set of protein encoding genes (Lauc and Zoldos, 2010). The plasticity of the glycoproteome is thought to allow higher eukaryotes to adapt and respond to selective pressures imposed by environmental factors or pathogens with faster generation times (Lauc and Zoldos, 2010; Royle et al., 2008; Varki, 1993).

1.1.4 Biosynthesis of Protein Glycans

Glycosylation primarily occurs in the endoplasmic reticulum (ER) and Golgi apparatus by the highly regulated action of glycosylation enzymes; glycosyltransferases transfer specific monosaccharides from uridine diphosphate (UDP) and guanosine diphosphate (GDP) complexes to glycan acceptors in a specific anomeric linkage (α or β); conversely, glycosidases remove specific monosaccharides (Figure 1.7) (Varki, 2009). Studies have shown that glycosyltransferases can be localised to specific subcellular compartments, which

can influence the final glycan structure (Colley, 1997). Glycosyltransferases also exhibit high specificity to certain glycan substrates in requiring the presence of a particular monosaccharide and linkage to function (Patenaude et al., 2002).

O-linked glycosylation most commonly involves the transfer of GalNAc to Ser or Thr residues; however, individual GlcNAc, Xyl, Man and Fuc residues can also be transferred (Varki, 2009). *O*-GalNAc glycosylation is initiated in the Golgi apparatus by one of 24 polypeptide-*N*-acetylgalactosaminetransferases (ppGalNacT-1 – 24). Each ppGalNAcT has a different substrate specificity when transferring GalNAc from UDP-GalNAc to Ser or Thr residues resulting in highly controlled *O*-linked glycosylation and significant heterogeneity (Brockhausen, 1999). Elongation of *O*-GalNAc is catalysed by glycosyltransferases, which attach monosaccharides one at a time to generate eight *O*-linked glycan cores (Figure 1.5) followed by further elongation for more complex *O*-linked glycans.

N-linked glycosylation is a different pathway initiated on the cytosolic side of ER by the synthesis of the *N*-linked glycan precursor ($Glc_3Man_9GlcNAc_2$), which is then transferred to an Asn residue of a protein in the Asn-X-Ser/Thr consensus sequence by the oligosaccharyltransferase (OST) complex (Figure 1.7)(Hubbard and Ivatt, 1981; Varki, 2009). A function of *N*-linked glycosylation is to ensure the proper folding of most membrane and secreted proteins by the calnexin (CNX) and calreticulin (CRT) pathway (Stronge et al., 2001). CNX and CRT bind to incompletely folded proteins that still contain at least one glucose residue on the *N*-linked glycan precursor, preventing their exit from the ER and enhancing folding efficiency.
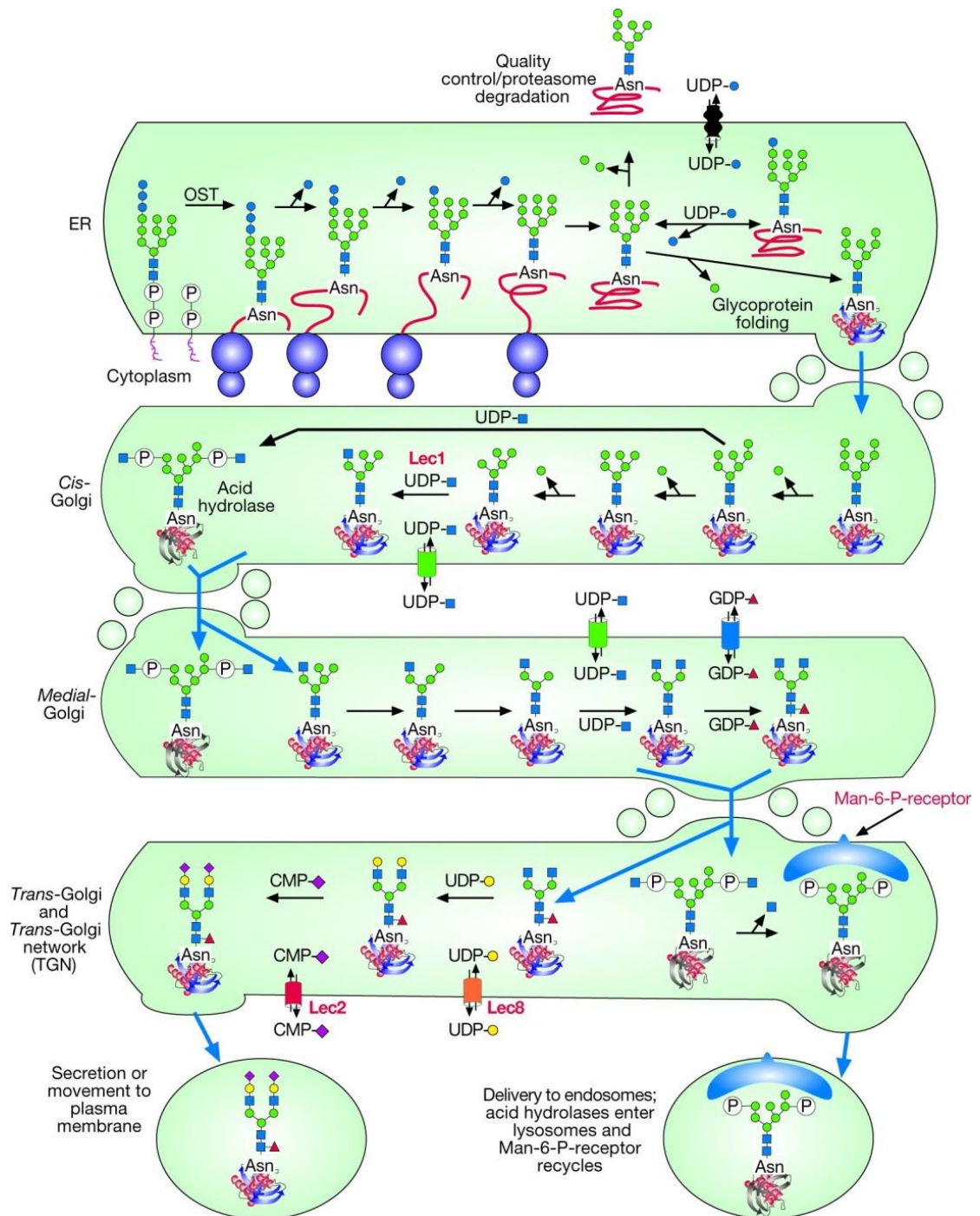
Figure 1.7 – Processing of *N*-linked glycans in the ER and Golgi apparatus. Reprint from open access Essentials of Glycobiology (2nd edition), Chapter 8, 2009 (Varki, 2009).

Once the protein is properly folded, the *N*-linked glycan precursor is processed by the sequential removal of glucose residues by α-glucosidases (I and II) and a single Man residue by α-mannosidase I, which then allows the protein to exit the ER and be translocated to the *cis*-Golgi (Figure 1.7)(Hubbard and Ivatt, 1981). Processing in the *cis*-Golgi involves the trimming of Man residues by $\alpha 1 - 2$ mannosidases (IA, IB, and IC) to produce $Man_5GlcNAc_2$,

a key intermediate in the synthesis of hybrid and complex *N*-linked glycans (Figure 1.6). Alternatively, trimming of Man can be limited or completely avoided to generate the high mannose class of *N*-linked glycans (Figure 1.6).

Biosynthesis of hybrid and complex *N*-linked glycans is initiated in the *medial*-Golgi; GlcNAc is transferred to the core of the $Man_5GlcNAc_2$ intermediate on the 3-arm by *N*-acetylglucosaminyltransferase I (GlcNAcT-1) via β1 − 4 linkage to yield the precursor of hybrid *N*-linked glycans (Figure 1.6) (Hubbard and Ivatt, 1981). Alternatively, the remaining Man residues on the 6-arm are trimmed by α-mannosidase II and GlcNAc is transferred to the 6-arm by GlcNAcT-2 via β1 − 4 linkage to yield the precursor of complex *N*-linked glycans (Figure 1.6)(Hubbard and Ivatt, 1981). Further elongation of hybrid and complex *N*-linked glycans can transfer bisecting or branching GlcNAc residues (Figure 1.8), Gal to terminal GlcNAc (Figure 1.9), Fuc to the core or terminal residues (Figure 1.10), NeuAc and other modifications such as sulphate, resulting in myriad number of possible combinations.
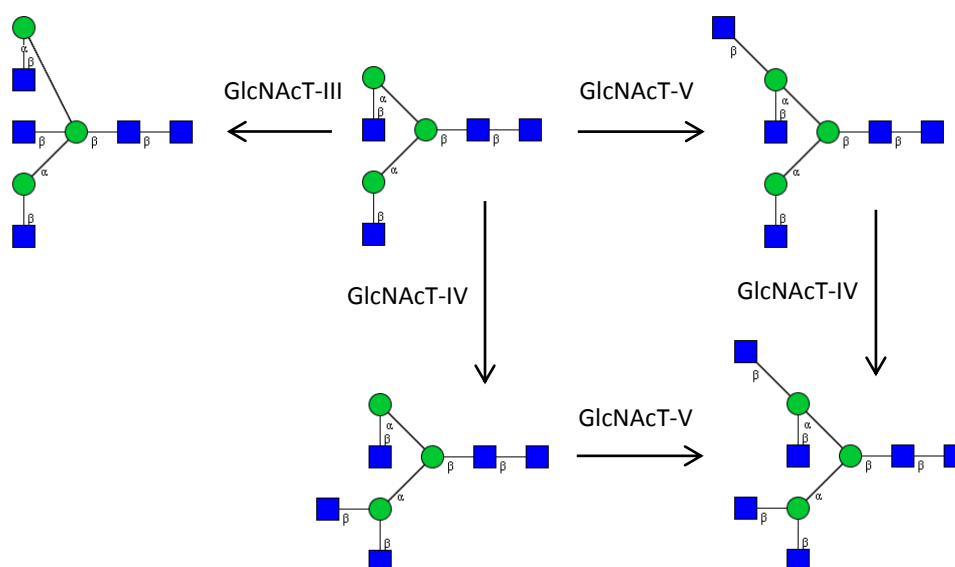
Figure 1.8 – Elongation of *N*-linked glycans by β-linkage of GlcNAc to core mannose residues. Linkage position is determined by enzyme substrate specificity. GlcNAcT-III transfers a bisecting GlcNAc while GlcNAcT-IV and GlcNAcT-V generate tri- and tetra-antennary structures.

Certain structures and configurations are more prevalent due to the substrate specificity of glycosyltransferases; elongation of hybrid and complex *N*-linked glycan structures most commonly occurs by the addition of galactose via β1 – 4 linkage to GlcNAc by β-1,4-galactosyltransferase to produce *N*-acetyllactosamine (Type 2 LacNAc); the addition of Gal via β1 – 3 linkage by β-1,3-galactosyltransferase is also possible but less common (Type 1 LacNAc) (Figure 1.9).
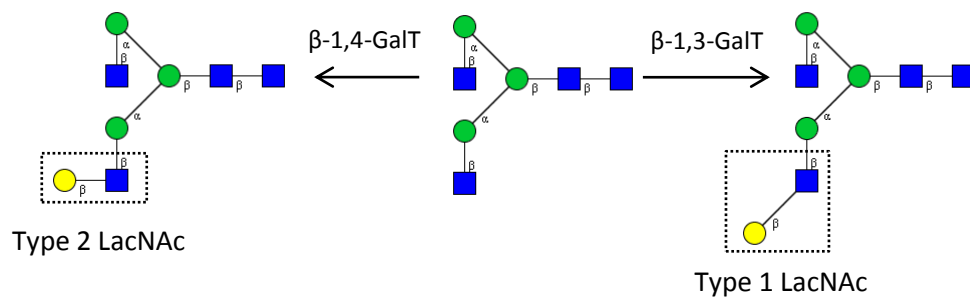
Figure 1.9 – Elongation of *N*-linked glycans by β1-4 and β1-3 linkage of galactose to GlcNAc to produce Type 2 LacNAc and Type 1 LacNAc respectively. Type 2 LacNAc is more common.

Fucose substitution always occurs via α linkage and can be transferred by fucosyltransferases to the *N*-linked glycan core via α1 – 6 linkage to the reducing end GlcNAc (in mammals), or terminal residues to produce the Lewis a/b and x/y epitopes by the substitution of terminal GlcNAc residues via α1 – 3 or α1 – 4 linkage depending on the position of Gal, or terminal Gal residues via α1 – 2 or α1 – 6 linkage (Figure 1.10)(de Vries et al., 2001; Kornfeld and Kornfeld, 1985).
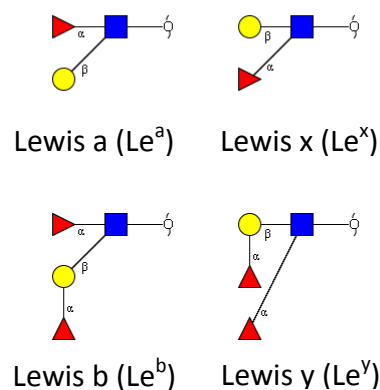
Figure 1.10 – The Lewis a/b and x/y epitopes as a result of fucose substitution.

An understanding of glycosyltransferase substrate specificities thus allows certain biosynthetic pathways to be assumed allowing the prediction of glycan structures, which is extremely useful in the analysis and structural characterisation of glycans.

1.2 Methods in Glycomics

As a subset of glycobiology, glycomics involves the systematic analysis of the "glycome", which encompasses all types of glycosylation expressed by a cell, tissue or organism and seeks to understand the relationship between glycosylation and a particular biological process (Rakus and Mahal, 2011; Varki, 2009). However, the nature and structural complexity (as described in Section 1.1.1) of glycans has presented unique challenges to the development of techniques for glycan analysis (Marino et al., 2010). As glycan biosynthesis is non-template driven, genetic-based techniques cannot provide a simple method of glycosylation prediction or analysis (Pratt and Bertozzi, 2005). Instead, glyco-profiling techniques have focused on detailed structural characterisation with ambitions to understand the role of glycan structure in disease states, discover potential biomarkers for disease diagnosis and prognosis, and to produce novel glycan based therapeutics (Packer et al., 2008). There is very little information in the scientific literature on preparative scale isolation of pure *N*-linked glycans from glycoproteins. The overwhelming majority of work within the field has instead focussed on the analytical scale isolation of *N*-linked glycans from very minute sample sizes. This thesis concentrates on the glycoprotein component of the glycome, so methods currently used for glycan release from proteins, and the separation and structural characterisation of the released glycans are described below.

1.2.1 The Release of Protein Glycans

Glycans that are conjugated to proteins can be released via enzymatic or chemical means resulting in reducing or reduced glycans respectively. A reducing glycan refers to an *N*- or *O*-linked glycan that contains the intact anomeric ring from carbon that was attached to the protein and is therefore capable of further reactions. Conversely, a reduced glycan refers to a released *N*- or *O*-linked glycan with the anomeric carbon converted to linear vicinal diols as a result of a reduction reaction.

Chemical release can be achieved by hydrazinolysis using anhydrous hydrazine (Takasaki et al., 1982) or alkaline β-elimination (Wu et al., 1984). By hydrazinolysis, both *N*- and *O*-linked glycans can be obtained separately by sequential release with *O*-linked glycans released first under mild conditions (60 °C) followed by *N*-linked glycan release under harsher conditions (90 °C). Though capable of obtaining reducing glycans at high yield, hydrazinolysis is rarely used in an academic setting due to the highly hazardous and reactive reagents and reaction conditions.

Alkaline β-elimination is commonly used to release *O*-linked glycans and utilises alkaline conditions to hydrolyse the glycosidic bond between glycans and hydroxyl groups of Ser and Thr residues (Marino et al., 2010). However, due to the high pH required for release that can lead to glycan degradation by isomerisation and subsequent "peeling", the reaction is usually performed in the presence of a reducing agent such as sodium borohydride which produces reduced glycans, thereby preventing such "peeling" by preventing isomerisation.

Enzymatic release is the current standard technique for the release of *N*-linked glycans from proteins and involves the use of amidases and endoglycosidases. The most widely used enzyme Peptide N-Glycosidase F (PNGase F), is actually an amidase that specifically cleaves *N*-linked glycans while leaving the reducing end intact and converts the asparagine residue in the *N*-glycosylation consensus sequence into aspartic acid (Maley et al., 1989; Tarentino et al., 1985). This has the advantage of a one mass unit increment on the deglycosylated peptide allowing identification of *N*-glycosylation sites and their occupancy by mass spectrometry. PNGase F cleaves all classes of *N*-linked glycans except those containing a core α1−3 substituted fucose, commonly found in plants and insects (Tretter et al., 1991). Release of core α1−3 fucose substituted *N*-linked glycans from glycopeptides is facilitated by another amidase, PNGase A (Plummer and Tarentino, 1981). Alternatively, enzymatic release by endoglycosidase D (Endo D) cleaves between the two GlcNAc residues of the chitobiose core

of all classes of *N*-linked glycans while endoglycosidase H (Endo H) cleaves in a similar fashion but only acts on high mannose and hybrid type structures (Figure 1.5) (Maley et al., 1989). Enzymatic release of *O*-linked glycans is limited to endo-α-*N*-acetylgalactosaminidase (*O*-glycanase), which is only active against the core 1 (Galβ1 – 3GalNAc-Ser/Thr) structure (Figure 1.4) (Maley et al., 1989).

1.2.2 Chromatographic Separation of Glycans

After release from the protein, separation of the glycans can be achieved by chromatography. Chromatography is a separation technique that relies on the differential affinity of analytes for a liquid phase as it migrates through a stationary phase (Lough and Wainer, 1995). Chromatographic techniques are defined by the chemistry of the stationary phase employed, resulting in affinities for specific types of molecules. Standard reversed phase (RP) chromatography used routinely in proteomics employs a non-polar stationary phase comprised of particles decorated with chains of eighteen carbons (C18). RP chromatography is thus not compatible with free glycans due to their polar nature and requires prior derivatisation of oligosaccharides with a hydrophobic agent such as 2-aminobenzamide (2-AB) for chromatographic separation.

The development of stationary phases suited for the analysis of underivatised glycans has resulted in two major techniques now being used for the separation of glycans: porous graphitised carbon (PGC) and hydrophilic interaction liquid chromatography (HILIC).

The PGC stationary phase is comprised of fully porous spherical graphite particles arranged in hexagonal sheets (Pereira, 2008). Though the exact mechanism is yet unknown, separation is described to occur via the polar retention effect on graphite (PREG), which involves polar interactions and is influenced by planarity, the presence of functional groups and charge (Pereira, 2008). The use of PGC-LC coupled to mass spectrometry is an established method of glycan analysis capable of separating isomeric structures (Jensen et al., 2012; Pabst et al., 2007; Ruhaak et al., 2009). Both reducing and reduced glycans are strongly retained by PGC

which are loaded in aqueous solvent and can be separated by gradient elution with organic solvent (usually acetonitrile) with a volatile additive (e.g. formic acid, trifluroacetic acid, ammonia, ammonium bicarbonate) (Ruhaak et al., 2009).

HILIC separation of glycans is performed using a range of hydrophilic stationary phases including bare silica, amine-, hydroxyl-, amide-bonded or zwitterionic particles (Zauner et al., 2011). Recent advances in HILIC have focused on high-throughput applications and new stationary phases that are capable of separating isomeric oligosaccharide structures. The separation of polar glycans is determined by hydrophilic potential and involves partitioning between the mobile phase and the hydrated surface of the stationary phase as a result of hydrogen bonding, ion exchange and dipole-dipole interactions (Zauner et al., 2011). Glycans are loaded in high organic solvent (~75%) and eluted by increasing the concentration of aqueous solvent with low concentrations of volatile acid or salt (<100 mM). PGC was selected for used in this project considering that PGC is most widely applied to underivatised glycans while HILIC is more suited to glycans derivatised with labels such as 2-AB (Ruhaak et al., 2010). Labelling reactions can be applied directly to the HILIC column without prior cleanup since glycans are loaded in high organic solvent, in which excess label is eluted while glycans are retained on the column

Other chromatographic techniques such as high performance anion exchange and affinity chromatography are also available In high performance anion exchange chromatography (HPAEC), glycans are deprotonated at high pH then separated on ion exchange resins (Kandzia and Costa, 2013). Retention on HPAEC is determined by the accessibility of the ionised glycan to the active site of the stationary phase (Kandzia and Costa, 2013). As glycans have little UV absorbance, HPAEC is typically coupled with pulsed amperometric detection (PAD) for detection which relies on the electrochemical responses of sugar ions in solution (Kandzia and Costa, 2013). However, HPAEC-PAD is generally not compatible with

downstream mass spectrometric (MS) analysis due to the high pH and amount of salts in the liquid phase.

1.2.3 Structural Characterisation of Glycans

Mass spectrometry (MS) is an established technique used for glycan characterisation offering high sensitivity and accuracy, low sample consumption and the ability to generate important structural information (Zaia, 2008). Mass spectrometry measures the mass of an ionised molecule and provides information in a mass spectrum that relates a mass-to-charge ratio (*m/z*) of a molecule to its abundance. A mass spectrometer is fundamentally comprised of three parts: an ion source, which is the region where ions are formed, a mass analyser where ions are separated according to their *m/z*, and a detector that records and converts signals into a mass spectrum (Herbert and Johnstone, 2010). A variety of ionisation techniques and mass analysers exist. Ionisation can be performed in positive and negative polarity, which influences the ionisation efficiency and sensitivity in the analysis of neutral and acidic glycans (Herbert and Johnstone, 2010).

In the process of matrix assisted laser desorption/ionisation (MALDI), an analyte is crystallised with an excess (1000 fold) of a low molecular weight ultraviolet (UV) absorbing matrix in the ion source. A laser is then directed continuously at the matrix which rapidly accumulates energy resulting in vaporisation of both matrix and analyte into a dense plume of gaseous ions (Wilm, 2011). The precise mechanism is yet unknown, however, ionisation is thought to involve transfer of energy from matrix ions to the analyte ions causing desorption and ionisation. MALDI is usually linked to a time-of-flight (TOF) mass analyser that separates ions by time and mass dependent ion velocities; heavy ions fly for prolonged periods with slower velocities while lighter ions fly for brief periods with faster velocities (Herbert and Johnstone, 2010).

In electrospray ionisation, molecules are ionised directly from the liquid phase by a spray atomiser in an electric field which produces a fine spray of droplets that undergo rapid evaporation and begin to eject solvated ions (Wilm, 2011). The ESI mass spectrometer used within the research group is equipped with an ion trap mass analyser which uses electrical and magnetic fields to trap and sort ions by their *m/z* (Kelley et al., 1988). Analysis by ESI with an ion trap is a continuous scanning event that allows the selection of high abundant ions for further fragmentation as an aid to molecular structural characterisation (discussed further in Chapter 5). Coupling of liquid chromatography and ESI mass spectrometry (LC-ESI-MS) is considered an online technique that combines separation and structural analysis allowing more detailed analysis of individual structures. Conversely, techniques such as MALDI-TOF are not directly coupled to LC.

In addition to MS, nuclear magnetic resonance (NMR) spectroscopy has proven to be a useful technique in elucidating tertiary fine structure such as anomeric linkage. Fundamentally, NMR relies on the magnetic properties of atomic nuclei in response to a magnetic field (Günther, 2013). A major limitation of NMR is the requirement of large amounts (~1 mg) of purified sample that are currently not realistically available or accessible from biological samples, which currently limits the application of NMR to synthetically produced glycans (Marino et al., 2010).

1.2.4 Purified Glycans as Reference Standards

The advancement of glycomics research partly relies on the availability of glycan standards which are essential for the development and validation of experimental techniques such as mass spectrometry and liquid chromatography. Glycan standards can be used in the calibration of mass spectrometers, provide a tool for the characterisation and quantitation of unknown glycans in a sample. By spiking well characterised glycan standards into mixtures, a library of standardised retention times can be generated allowing structural identification with

the added dimension of reproducible relative retention times (Pabst et al., 2007). Possessing a glycan standard reference also enables direct comparison between the efficacies of different modes of analysis, a process that is essential for the optimisation of novel techniques (Harvey et al., 2008).

Glycan standards are in high demand not only for their many applications but also due to the difficulty in acquiring adequate quantities of sufficient purity. Despite recent advances in solid phase synthesis, the *de novo* chemical synthesis of complex glycans is limited by low stereoselectivity, low overall yields, lengthy production time, and often require the use of highly toxic reagents (Han et al., 2012; Lepenies et al., 2010).

While a range of glycan standards are commercially available from a growing number of companies, the costs are considerable which limits frequent use and availability to small scale limited experiments. Depending on structure, a typical standard costs between \$300 – \$500 per 10 micrograms. A search of suppliers found that high mannose and bi-, tri-, and tetra-antennary *N*-linked glycans were the most common types of standards available (Table 1.1).

Table 1.1 – Companies producing commercially available glycan standards. Standards are available as the asialo (no NeuAc) glycoform or with varying degrees of sialylation in the NeuAc column in both α2 -3 and α2 – 6 linkages. Information obtained from respective online catalogues; Dextra: http://www.dextrauk.com; Prozyme: http://www.prozyme.com; QA-Bio: http://glycotools.qa-bio.com/glycans; Sigma Aldrich: http://www.sigmaaldrich.com.

Company	*N*-linked glycan Standard	NeuAc α2 – 3,6	Protein Source
Dextra	Tetra-antennary		
	Tri-antennary	3	
	Bi-antennary	1,2	Sources not provided
	Bi-antennary + core fucose	1,2	
	Bi-antennary bisecting GlcNAc + core fucose		
	High mannose (Man1 – Man9)		
Prozyme	Tetra-antennary		Human a1-acid glycoprotein
	Tri-antennary	3	Bovine fetuin
	Bi-antennary	1,2	Human fibrinogen
	Bi-antennary + core fucose	1,2	Porcine thyroglobulin
	Bi-antennary bisecting GlcNAc + core fucose		Porcine thyroglobulin
	High mannose (Man5 – Man9)		Porcine thyroglobulin
QA-Bio	Tetra-antennary		Human a1-acid glycoprotein
	Tri-antennary	3	Bovine serum fetuin
	Bi-antennary	1,2	Fibrin
	Bi-antennary + core fucose	1,2	Porcine thyroglobulin
	High mannose (Man1 – Man9)		Porcine thyroglobulin
Sigma Aldrich	Tri-antennary	3	Bovine (protein not specified)
	Bi-antennary		Human fibrinogen
	Bi-antennary + core fucose	1,2	Porcine thyroglobulin
	High mannose (Man9)		Porcine thyroglobulin

Bi-antennary *N*-linked glycans containing a core substituted fucose and a bisecting GlcNAc (core fucosylated form only) were also available. However, there is a distinct lack of standards containing terminally substituted fucose such as those in the lewis a/x and b/y epitopes, which have been implicated in the pathogenesis of bacterial species such as *E. coli* and are of increasing interest for the development of oligosaccharide based therapeutics (Le Pendu, 2004; Newburg et al., 1990; Newburg et al., 2004).

Human milk proteins have been reported to express a significant degree of fucosylation in comparison to bovine milk proteins (Dallas et al., 2011; Nwosu et al., 2012). As such, human milk was identified as a good source of terminally fucosylated *N*-linked glycans that could potentially fill this lacking niche of *N*-linked glycan standards as well as providing the more common structures in relatively high concentrations. Human milk glycosylation is described in greater detail in Chapter 3.

1.3 Project Aims

The major aim of this project was to develop a protocol to release and purify N-linked glycans from proteins on a preparative scale, i.e. in quantities suitable for downstream applications as glycan standards. Human milk was selected as a suitable source of N-linked glycans for this project since milk proteins are known to be highly glycosylated and human N-linked glycans are of particular interest for medical research due to growing evidence of their involvement in the intercellular and host-pathogen interactions. Furthermore, human milk glycoproteins exhibit a high degree of fucosylation, a substitution that is underrepresented in the current range of commercially available glycan standards.

The specific aims of this project were to:

1. Develop a method for preparative scale release of N-linked glycans from human milk proteins (Chapter 3)

2. Optimise a method for the preparative scale separation and offline detection of released N-linked glycans in order to obtain purified standards (Chapter 4)

3. Establish a PGC-LC retention time library and ESI-MS fragmentation dataset using the purified standards (Chapter 5)

4. Test the suitability of the glycan standards for application in a biological assay, i.e. to investigate whether the N-linked glycans released from human milk proteins could act as an inhibitor of bacterial binding to cells (Chapter 6)

Chapter 2 – General Materials and Methods

2.1 Materials

Materials and reagents used in this work are shown in Table 2.1. All other materials and reagents were purchased from Sigma Aldrich unless stated otherwise. MilliQ® ultrapure water was collected from a MilliQ Synthesis system (Millipore).

Table 2.1 – Materials and reagents used in this project.

Reagent	Supplier	Used in
Acrowell 96 Filter plate 0.45 µg Biotrace PVDF	Pall Corporation	Section 6.2.3
250 mL Cell Culture flask	Thermo Scientific	Section 6.2.1
100 mL Cell Culture plate	Thermo Scientific	Section 6.2.1
Dowex AG-50W-X8 cation exchange resin 300–1,180 µm wet bead size	Biorad	Section 4.2.1
Glyko® α(1-3,4)-Fucosidase (almond meal)	Prozyme	Section 2.2.6
Glyko® β(1-4)-Galactosidase (*Streptococcus pneumoniae*)	Prozyme	Section 2.2.6
Reaction Buffer (250 m*M* sodium acetate, pH 5.0) 5X	Prozyme	Section 2.2.6
HyperSep Hypercarb graphitised carbon solid phase extraction (SPE) column	Thermo Scientific	Section 4.2.1
Immobilon-P™ Transfer Membrane	Millipore	Section 2.2.4
Mammalian protease inhibitor cocktail	Novagen	Section 3.2.1
Novex Sharp pre-stained protein standard	Life Technologies	Section 2.2.1
NuPAGE 4-12% Bis-Tris 1.0 mm x 12 well gels	Life Technologies	Section 2.2.1
NuPAGE LDS Sample buffer 4X	Life Technologies	Section 2.2.1
NuPAGE MOPS Running Buffer 20X	Life Technologies	Section 2.2.1
Peptide N glycosidase F	Roche	Section 2.2.4, Section 3.2.4, Section 3.2.6
Polypropylene 96 well plate	Corning Incorporated	Section 2.2.4
RPMI 1640 media 1X	Gibco	Section 6.2.1
Strata X Polymeric Reversed Phase Cartridge	Phenomenex	Section 4.2.1
SYBR Green I	Life Technologies	Section 6.2.3
Tryptone	Oxoid	Section 4.2.2
Zeba© desalt spin column 2mL	Thermo Scientific	Section 3.2.6

2.2 General methods

2.2.1 Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE) was performed with NuPAGE 4 – 12% Bis-Tris 1.0 mm x 12 well gels. Samples were prepared by adding 10 µL NuPAGE lithium dodecyl sulphate (LDS) sample buffer (4X) to 10 µL of sample, 10 µL MilliQ water and 1 µL 1M dithiolthreitol (DTT) followed by heating at 90 °C for 5 mins. Equal amounts (if possible) of protein were loaded in parallel with Novex Sharp pre-stained protein standard and electrophoresed at 200V for 50 mins in a running buffer containing 50 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 50 mM Tris base, 0.1% SDS, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.7, prepared by mixing 50 mL NuPAGE MOPS Running buffer (20X) with 950 mL MilliQ water.

After electrophoresis, gels were stained with Coomassie blue protein stain prepared by mixing 40 mL of Solution A (3 g Coomassie blue G-250, 400 mL methanol, 47 mL 85% phosphoric acid, 1165 mL MilliQ water) with 10 mL of Solution B (250 g ammonium sulphate, 500 mL MilliQ water). Gels were stained overnight on a rocker and de-stained in 1% acetic acid.

2.2.2 Bradford Assay for Protein Quantification

The Bradford assay for protein quantification was performed in a 96 well plate format. Proteins standards of bovine serum albumin (BSA) were prepared to concentrations of 0.1 mg/mL, 0.25 mg/mL, 0.5 mg/mL, 1.0 mg/mL, and 1.4 mg/mL. In triplicate, 5 µL of sample and 5 µL of each BSA standard were added to separate wells. To each well, 250 µL of Bradford Reagent was then added and gently mixed on a shaker for 1 min. Samples were allowed to incubate for 10 mins before measurement of the UV absorbance at 595 nm on a Fluostar plate reader (BMG technologies). A standard curve was obtained by plotting the UV absorbance against the BSA standard protein concentrations. The unknown sample concentration was then determined by comparing the absorbance against the standard curve.

2.2.3 Periodic Acid – Schiff's Reagent (PAS) Assay for Glycan Quantification

The PAS stain is a colourimetric stain commonly used to visualise glycoproteins and polysaccharides. The PAS stain was performed in a 96 well plate format developed by Kilcoyne et al. (2011). Bovine fetuin was used to generate standard curve and solutions were prepared to concentrations of 0.25 mg/mL, 0.5 mg/mL, 1.0 mg/mL, 2.0 mg/mL, and 5 mg/mL. To each well, 25 µL of sample and 25 µL of each fetuin standard were added in triplicate. 120 µL of periodic acid (0.06% in 7% acetic acid) was added and incubated at 37 °C for 90 mins. After incubation, 100 µL of Schiff's reagent was added and left for 45 mins for complete colour formation before measurement by UV absorbance at 550 nm on a Fluostar plate reader (BMG technologies). A standard curve was obtained by plotting the absorbance against the molar concentration of fetuin protein standard, factoring in the number of glycosylation sites. An approximate molar concentration of glycosylation was determined by plotting the absorbance of the unknown sample against the standard curve.

2.2.4 Analytical Scale "Dot blot" Glycan Release

Dot blotting to PVDF membrane

N- and *O*-linked glycans were released from proteins on an analytical scale according to Jensen et al. (2012) Proteins were immobilised onto Immobilon-P™ Transfer polyvinylidine fluoride (PVDF) membrane. The membrane was first wet with ethanol and proteins were dot-blotted when the surface of the membrane was dry. Once the protein was immobilised, the membrane was allowed to completely dry (between 1 hr and overnight depending on time constraints) to ensure all protein was properly bound then rewet by soaking in methanol for 15 minutes and washed with water for 15 mins to remove salts. The membrane was then stained with Direct Blue 71, prepared by mixing 1.6 mL Direct Blue 71 (0.1% stock solution) with wash solution (40% ethanol, 10% acetic acid). Once spots were visualised, the membrane was briefly de-stained in wash solution, washed with water and dried.

Glycan release with PNGase F

The remainder of the procedure was performed in polypropylene 96 well plates. Each well was first blocked with 100 µL 1% (w/w) polyvinyl pyrrolidone 40 000 (PVP) solution to avoid non-specific binding of PNGase F to the well. Once the PVP solution was removed, proteins spots were excised from the membrane and transferred to wells. PNGase F (5 µL, 0.5 Units/µL) was added to protein spots and incubated at 37 °C overnight. Wells were washed twice with 20 µL of MilliQ water to recover released glycans

Deamination and Reduction of N-linked glycans

Released *N*-linked glycans in a volume of 50 µL were deaminated by the addition of 10 µL of 100 mM ammonium acetate pH 5 and incubating at room temperature for 1 hr before being dried by vacuum centrifuge. Dried deaminated *N*-linked glycans were then reduced by the addition of 20 µL of 1M $NaBH_4$ in 50 mM KOH at 50 °C for 3 hrs. The reaction was quenched by the addition of 2 µL glacial acetic acid and desalted.

Desalting the reduced N-linked glycans

Mini ion exchange columns were prepared by applying ~30 µL bed volume Dowex AG-50W-X8 cation exchange resin onto a Ziptip® (Millipore). Columns were conditioned with 50 µL of 1M hydrochloric acid (HCl), 50 µL of methanol and 50 µL of MilliQ water three times for each step. Washing was performed using a tabletop mini-centrifuge to pull solutions through the columns. The reduced *N*-linked glycans were then applied to the mini-columns and eluted with 50 µL of MilliQ water (twice) then dried by vacuum centrifuge. Once dried, the *N*-linked glycans were dissolved in 200 µL of methanol to produce the volatile trimethylborate to remove residual borate by vacuum centrifugation. The dried reduced *N*-linked glycans were dissolved in 50 µL of MilliQ water and further purified by solid phase extraction (SPE) on graphitised carbon mini-column, prepared as above except with graphitised carbon from

HyperSep Hypercarb columns instead of Dowex. Columns were conditioned with 50 µL of 80% (v/v) acetonitrile, 0.05% (v/v) trifluoroacetic acid and 50 µL of MilliQ water three times. Reduced *N*-linked glycans were applied to the column and washed with 50 µL MilliQ water three times to remove salts. Reduced *N*-linked glycans were finally eluted with 40% (v/v) acetonitrile, 0.05% (v/v) trifluoroacteic acid.

2.2.5 Analysis of Glycans by Porous Graphitized Carbon Liquid Chromatography – Electrospray Ionisation Tandem Mass Spectrometry (PGC-LC-ESI-MS/MS)

PGC-LC-ESI-MS analysis was performed using an established method (Jensen et al., 2012). Reduced *N*-linked glycans (Section 2.2.4) were dissolved in 10 mM ammonium bicarbonate and separated by online liquid chromatography (LC) using a Hypercarb PGC 5 micron particle size 0.18 mm i.d. x 100 mm column maintained at 27 °C. The samples were loaded in 10 mM NH_4HCO_3, pH 8. Separations of *N*-linked glycans were performed over an 85 min gradient of 0 % - 45% (v/v) acetonitrile in 10 mM ammonium bicarbonate. The flow rate was maintained at 2 µL/min using a HPLC system (Agilent 1100, Agilent Technologies Inc., CA, USA) and the eluate was introduced directly into the ESI source (Agilent 6330, Agilent Technologies Inc., CA, USA). In the ion-trap MS, the voltage of the capillary outlet was set at 3kV, and the temperature of the transfer capillary was maintained at 300 °C. The MS spectra were obtained in the negative mode over a mass ranged between *m/z* 200 and 2200. Ions were detected in ion charge control (ICC) (target 80,000 ions) with an accumulation time of 200 ms. Induced collision was performed at 35% normalized collision energy and an isolation window of 4 *m/z*.

Chapter 3 – Preparative Scale Release of *N*-linked Glycans from Human Skim Milk Proteins

3.1 Introduction

The aim of the work described in this chapter was to develop a method of releasing *N*-linked glycans from human milk proteins on a preparative scale to potentially enable their subsequent use as structurally defined glycan standards. The *N*-linked glycans were of particular interest as they are in high demand as standards, being the most well studied class of glycan, and because synthetically produced *N*-linked glycan standards are difficult to acquire due to their large size and complexity (Rakus and Mahal, 2011). Human milk was selected as a good source of *N*-linked glycans partly due to its high protein content of which over 70% are glycosylated (Apweiler et al., 1999). In addition, the *in vitro* application of the released *N*-linked glycans could be trialled in a later phase of the project using an established in-house bacterial binding assay to investigate the contribution of human milk *N*-linked glycans to the anti-infective qualities of human milk (Chapter 6).

3.1.1 The Composition of Human Milk

Human milk is a complex mixture of macro- and micro-nutrients which can be separated into skim milk and cream fractions. The skim milk fraction contains water, carbohydrates in the form of lactose and human milk oligosaccharides (HMOs), proteins, and vitamins, and salts (Picciano, 2001). The cream fraction is mainly comprised of lipids surrounded by the milk fat globule membrane (MFGM) (Picciano, 2001). When reporting on the composition of human milk, exact concentrations of each macronutrient are somewhat arbitrary as there is significant fluctuation in the concentration and volume of milk over the course of lactation (Coppa et al., 1993). Variation in nutrient concentration between individuals is further influenced by factors including genetics, geography, diet and even time of day (Chaturvedi et al., 2001; Jensen, 1995; Macias and Schweigert, 2001; Saint et al., 1984; Shi et al., 2011; Wojcik et al., 2009;

Yamawaki et al., 2005). Human milk composition was reviewed in detail by Picciano M.F. (2001) and the following macronutrient concentrations were provided (Table 3.1).

Table 3.1– Approximate concentration of the macronutrients found in human milk.

Macronutrient	Concentration	Reference
Protein	7 – 8 g/L	(Lönnerdal, 2003)
Total lipids	30 – 40 g/L	(Jensen, 1999)
Human milk oligosaccharides (HMOs)	9 – 12 g/L	(Coppa et al., 1993)
Lactose	68 – 70 g/L	(Jensen, 1995)

The carbohydrate lactose is by far the most abundant component of human milk (68 – 70 g/L, Table 3.1). Though lactose is a disaccharide of galactose β1-4 linked to glucose, it is not considered an oligosaccharide because of its small size. HMOs comprise the third largest component of human milk (9 – 12 g/L, Table 3.1) and have been comprehensively studied as a unique class of oligosaccharide (Bode, 2012). All HMOs contain lactose at their reducing end and are elongated by the addition of lacoto-*N*-biose or *N*-acetyllactoasmine by β1–3 or β1–6 linkage. The lactose core or oligosaccharide chain can be further substituted with fucose or *N*-acetylneuraminic acid to produce a considerably diverse range of structures (Bode, 2012). While lactose and the HMOs are major carbohydrate components of human milk, both lack conjugation to proteins and so did not represent a source of the *N*-linked protein glycans that were the focus of this project.

The proteins found in human milk are diverse in function, involved in growth and maintenance, immunity support, and lipid, protein, and nucleic acid metabolism (Liao et al., 2011). Over 95% of human milk proteins are found in the skim milk fraction divided between the soluble whey proteins (~60%) and the insoluble caseins (~40%) (Rudloff and Kunz, 1997). The higher ratio of whey protein to casein is unique to human milk; caseins are significantly more abundant in the milk of other mammals, approaching 20:80 (whey:casein)

in bovine milk (Rudloff and Kunz, 1997). The remainder of milk proteins (~5%) are found associated with the MFGM, which is produced during the secretion of lipid globules which become enveloped in the apical plasma membrane of the alveolar cells of the mammary glands.

3.1.2 Human Milk Glycoconjugates

There are two major classes of macromolecules that are conjugated with glycans in human milk – the glycolipids and the glycoproteins. Glycolipids, or gangliosides, are almost exclusively found embedded within the MFGM. Glycolipids consist of a glycan epitope attached to a hydrophobic ceramide base which is anchored within the hydrophobic environment of the lipid globule while exposing the glycan component to the hydrophilic environment of milk (Jensen, 1999). Glycolipids occur at approximately 11 mg/L and as such are an extremely minor fraction of the total lipids found in mature human milk (30 – 40 g/L, Table 3.1) (Lægreid and Fuglesang, 1986).

Human milk glycoproteins are found in both the skim milk and cream fractions. While there have been numerous studies on individual glycoproteins such as lactoferrin, α-lactalbumin, secretory immunoglobulin A (sIgA), and lysozyme, very few studies have reported on the total glycosylation profile of human milk proteins (Peterson et al., 2013). In terms of glycosylation type, the whey and MFGM proteins exhibit both *N-* and *O-*linked glycosylation while the caseins (only κ-casein) exhibit only *O-*linked glycosylation (Raiha, 1985).

Only two studies have reported on the total *N-*linked glycosylation of human milk whey proteins. Nwosu et al. (2012) compared the *N-*linked glycosylation of human and bovine milk whey proteins. Significant differences in the degree of fucosylation and sialylation were observed with approximately 75% of human milk whey glycoproteins containing fucosylation compared to just ~31% of bovine whey glycoproteins (Nwosu et al., 2012). In terms of sialylation, roughly 57% of human whey glycoproteins compared to ~68% of bovine whey

glycoproteins contained at least one sialic acid (Nwosu et al., 2012). An earlier study by Dallas *et al.* (2011) reported on the complete *N*-linked glycosylation profile of mature human milk. Of the total *N*-linked glycan compositions identified ~65% were fucosylated, 38% were sialylated and 25% contained both fucosylation and sialylation (Dallas et al., 2011). Both studies were performed using high-performance microfluidic chip liquid chromatography coupled to mass spectrometry and were optimised on an analytical scale. In addition to the "dot-blot" protocol developed in-house for the release of glycans (Jensen et al., 2012), the analytical scale procedures described by Nwosu *et al.* (2012) and Dallas *et al.* (2011) for the preparation of whey proteins and release of *N*-linked glycans were used as a basis for the preparative scale work reported here.

3.1.3 Preparative Scale Release of *N*-linked Glycans

Chemical and enzymatic reactions are the two main methods of releasing *N*-linked glycans from glycoproteins (Section 1). Preparative scale hydrazinolysis is most commonly used for large scale industrial releases of *N*-linked glycans from gram quantities of glycoproteins (Evers et al., 1998). However, the required reaction conditions can be difficult to reproduce in that the hydrazine cannot be allowed to contact oxygen, and the glycoprotein sample must be very dry to prevent hydrazinolysis of NeuAc (El-Rassi, 2002). In addition, large scale hydrazinolysis is extremely hazardous due to the potential ignition of hydrazine and complicated by extensive purification to remove excess hydrazine.

Enzymatic release offers an alternative to chemical release that takes advantage of the availability of endoglycosidases such as PNGase F (Stubbs et al., 1997). The main drawback of enzymatic release is cost as the reactions become prohibitively expensive during scale-up.

3.2 Development of a Preparative Scale Glycan Release from Human Milk Glycoproteins

In developing a preparative scale release of *N*-linked glycans from human milk, a number of factors were considered. These included separation of human milk proteins from other milk components, selection of a suitable method of glycan release, maximising the purity and recovery of glycans from reaction solutions, reproducibility, and practicality.

3.2.1 Preparation of Human Skim Milk

The skim milk fraction was selected for protein extraction as it was identified to contain the majority of *N*-linked glycosylated proteins within the whey proteins whilst the cream fraction accounts for just $2 - 4\%$ of the total protein in human milk (O'Donnell et al., 2004). Whole human milk was thawed and 1 µ/mL of mammalian protease inhibitor cocktail was added. Human skim milk was prepared by centrifuging whole human milk at 13, 000 rcf, 4 °C for 30 minutes in 50 mL Falcon tubes (VWR International) forming a thick layer of cream that floated above the skim milk. Removal of the cream by pipette or spatula was difficult as the cream readily mixed into the skim milk upon the slightest agitation. To simplify removal of the cream, the tubes were frozen at -20 °C and the solidified cream was removed with a spatula.

3.2.2 Precipitation of Human Skim Milk Proteins

In order to isolate proteins from skim milk, precipitation using trichloroacetic acid (TCA), ethanol, acetone, and chloroform/methanol were considered. Studies that have compared the effectiveness of different precipitation procedures found very little difference in the recovery of proteins (Jiang et al., 2004; Zellner et al., 2005). Since lactose and HMOs are composed of some of the same monosaccharides as protein linked glycans, it was important to remove lactose and HMOs from the skim milk fraction to ensure there was no cross contamination with the released *N*-linked glycans which could later cause analytical interference. The

selection of a precipitation procedure for this project was thus based on practicality and the ability to selectively precipitate the human skim proteins free from lactose and HMOs without damaging the protein linked glycans. The TCA precipitation method (Fountoulakis, 2001) was excluded since glycans containing N-acetylneuraminic acid (Neu5Ac) are highly susceptible to acid hydrolysis and the use of TCA, which is a strong acid, could potentially result in the loss of Neu5Ac from protein glycans during precipitation. Precipitation with acetone has been shown to precipitate lactose and HMOs in addition to proteins and was therefore excluded (Verostek et al., 2000). Ethanol precipitation reportedly maintains HMOs in solution (Stahl et al., 1994) but salts and vitamins are known to have low solubility in ethanol and can contaminate the precipitated proteins (Zellner et al., 2005). The chloroform/methanol precipitation method was considered an appropriate method for the precipitation of milk proteins as it takes advantage of the phase separation between aqueous and organic solvents (Wessel and Flügge, 1984).

Approach: To 1 volume of skim milk, 4 volumes of methanol, 2 volumes of chloroform and 3 volumes of water were added with vigorous mixing by vortex between each addition. The mixture was centrifuged at 10, 000 rcf, 4 °C for 30 minutes to collect proteins at the interphase between the aqueous and organic phases. The upper aqueous phase was removed and 3 volumes of methanol were added to the remaining interphase and organic phase. The mixture was centrifuged again at 10, 000 rcf, 4 °C for 30 minutes to pellet the proteins and the supernatant was removed. The pellet was then washed with methanol to remove traces of the organic phase, centrifuged again and the supernatant removed. While Wessel and Flügge reported a total precipitation volume of 9 mL of solvent to 1 mL of protein solution, the procedure was optimised in this work for 35 mL of solvent starting with 15 ml of skim milk.

Outcome: Chloroform/methanol precipitation was chosen as an appropriate method for precipitating the milk proteins (Section 3.2.1) since other milk components (lactose, HMOs,

vitamins, and salts) reportedly remain in the aqueous phase and residual lipids of the cream fraction dissolve into the organic phase (Wessel and Flügge, 1984). The aqueous phase on top the organic phase was easily removed by pipette. The addition of methanol to the organic phase caused a decrease in the density of the organic phase allowing the precipitated proteins at the interphase to be pelleted by centrifugation, which otherwise would remain afloat on top of the organic phase. Despite the advantages of the chloroform/methanol precipitation it was found that that the precipitated skim milk proteins were completely insoluble once dried. Attempts at redissolving the precipitate in a range of buffers including 8M urea and up to 10% SDS were unsuccessful. As the desired method of *N*-linked glycan release was enzymatic, it was imperative that the skim milk proteins be completely dissolved in preparation for enzymatic glycan release with PNGase F, so other methods of protein precipitation from milk were explored.

3.2.3 Removal of Casein by pH and Calcium Adjustment and Ultracentrifugation

Further investigation in the literature helped attribute the insolubility of the skim milk protein precipitate (Section 3.2.2) to the presence of caseins. Caseins are a family of phosphorylated proteins ubiquitous to milk and are responsible for its white cloudy appearance (Waugh and von Hippel, 1956). There are only two caseins in human milk which exist associated together in colloidal suspension as casein micelles. The major β-casein (~75% of casein) is hydrophobic and precipitates in the presence of calcium while the minor κ-casein (~25% of casein) is hydrophilic and contains multiple phosphorylated calcium binding sites (Fox and Brodkorb, 2008; Liu and Guo, 2008). Though there has been much discussion on the formation and stabilisation of the casein micelle, the exact structure of the micelle is still under debate (Dziuba and Minkiewicz, 1996; Fox and Brodkorb, 2008). However, the most widely accepted model places the hydrophobic β-casein localised within the core of the micelle that is surrounded by κ-casein, which functions to stabilise the micelle by acting as the interface between the hydrophobic core and the milk serum (Guo et al., 2003; Waugh and

von Hippel, 1956). It was possible that the precipitation of proteins with the chloroform/methanol method disrupted the structure of the micelle, liberating the hydrophobic β-casein and contributing to the overall insolubility of the skim milk protein precipitate. Upon review of the human caseins, it was found that β-casein is un-glycosylated and κ-casein contains only *O*-linked glycans (Ginger and Grigor, 1999; Madadlou et al., 2009). Since the aim of this work was to obtain released *N*-linked glycans, the caseins were considered redundant and the removal of caseins prior to organic solvent precipitation was explored in an attempt to improve the solubility of the skim milk protein precipitate.

A study by Kunz and Lönnerdal (1990) found that caseins could be precipitated from skim milk by adjusting the pH to 4.6 followed by ultra-centrifugation at 100 000 g for 90 minutes. The adjustment of pH was thought to disrupt the calcium binding interactions of κ-casein resulting in destabilisation of the casein micelle. When the pH was further reduced to 4.3, an increase in the precipitation of caseins was observed but also caused increased precipitation of whey proteins (Kunz and Lönnerdal, 1990). The same study also demonstrated that the addition of $CaCl_2$ greatly increased the precipitation of caseins and reduced contamination of the casein pellet with whey proteins. The rationale for calcium addition was to target the sensitivity of β-casein to calcium and encourage precipitation. Therefore, in this work, the method of Kunz and Lönnerdal (1990) was trialled in an attempt to remove casein from human skim milk with the aim of improving the solubility of the skim milk protein precipitate.

Approach: Caseins were precipitated from skim milk (Section 3.2.1) by pH adjustment to 4.6 with 1M HCl while on ice and the addition of 1M $CaCl_2$. Final concentrations of between 60 mM and 100 mM $CaCl_2$ were trialled with the most effective removal of casein achieved with 100mM $CaCl_2$. Acid adjustment and calcium precipitated caseins were pelleted then by ultracentrifugation at 100 000 g (Discovery Ultracentrifuge 90SE, Sorvall), 4°C for 90

minutes and the casein-depleted skim milk supernatant was collected. Proteins from the supernatant were then isolated by chloroform/methanol precipitation as above (Section 3.2.2).

Outcome: The extent of casein depletion was monitored by SDS-PAGE and it was clearly evident that caseins were successfully depleted from the skim milk by pH adjustment, calcium addition and ultra-centrifugation (Figure 3.1).

Human skim milk was included as a reference (Figure 3.1, lane 2) though more protein was present in lane 3 and lane 4 as can be seen in the Coomassie Blue stained gel (Figure 3.1, lane 3, lane 4). It is clear that the whey proteins were enriched at the expense of casein (Figure 3.1, lane 3). Proteins from the acid/calcium precipitated pellet (Figure 3.1, lane 4) formed a large distinct band in the molecular weight range of the caseins (κ-casein: 25 000 – 30 000, β-casein: ~24 000 (Kunz and Lönnerdal, 1990)).

Some other faint protein bands were visible suggesting that very few whey proteins were precipitated (Figure 3.1, lane 4). In the casein-depleted skim milk protein fraction (Figure 3.1, lane 3), protein bands corresponding to the expected MW of some abundant whey proteins could be seen including secretory immunoglobulin A (sIgA, MW: ~160 kDa (Liu and Newburg, 2013)), lactoferrin (MW: ~80 kDa (Liu and Newburg, 2013)), serum albumin (MW: ~67 kDa (Liu and Newburg, 2013)), lysozyme (MW: ~14 kDa (Liu and Newburg, 2013)) and α-lactalbumin (MW: ~14 kDa (Liu and Newburg, 2013)) and their expected molecular weights are indicated by arrows (Figure 3.1).
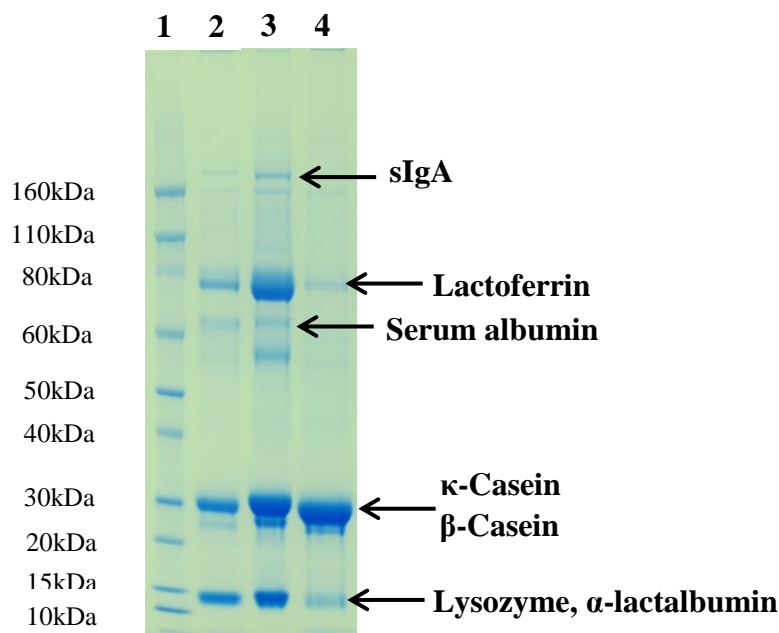
Figure 3.1 – Human milk protein fractions separated on an SDS-PAGE gel stained with Coomassie Blue. Lane 1: molecular weight markers; lane 2: skim milk (10 µL); lane 3: skim milk protein fraction following removal of casein by pH adjustment, calcium addition and ultracentrifugation (~10µg); lane 4: casein pellet (~10 µg). The expected molecular weights of major milk proteins are indicated by arrows (Liu and Newburg, 2013).

A distinct band was still observed around 30kDa indicating that some caseins were still present (Figure 3.1, lane 3). Nonetheless, the method of Kunz and Lönnerdal (1990) resulted in sufficient removal of casein to enable complete solubilisation of the remaining milk proteins (predominantly whey proteins) in 10 mM ammonium bicarbonate with 10% SDS.

3.2.4 Enzymatic In-solution Preparative Scale Release of *N*-linked Glycans

It is generally accepted that intact glycoproteins cannot be enzymatically de-glycosylated efficiently without first denaturing in the presence of detergent to increase accessibility of the enzyme to the glycosylation site (Tarentino et al., 1985). A detergent is a surfactant that is amphiphilic in that it contains both hydrophilic and hydrophobic groups (Zana, 1980). When used to solubilise and denature proteins, the hydrophobic tails of detergent molecules interact with the hydrophobic regions of proteins and disrupt intramolecular hydrophobic interactions. As the hydrophilic heads interact with the aqueous solution, the hydrophobic regions are able to interact with the aqueous solution and the structure of the protein dissociates and "denatures" (le Maire et al., 2000). Denaturing with dithiothreitol (DTT) and iodoacetamide

(IAA) are also commonly used with PNGase F digestion (Jensen et al., 2012). DTT is a reducing agent that breaks the disulfide bonds between cysteine residues that stabilise the structure of proteins (Jensen et al., 2012). Reduction with DTT is commonly followed by IAA which binds irreversibly to the reduced free thiol groups of cysteine residues preventing the reformation of disulfide bonds (Smythe, 1936).

While PNGase F does not contain disulfide bonds and is therefore unaffected by DTT and IAA, SDS and other ionic detergents are known to strongly inhibit PNGase F activity (Maley et al., 1989). Despite this, Nwosu et al. (2012) reported the successful release of *N*-linked glycans from human milk proteins using PNGase F in the presence of SDS. Whey proteins were first isolated from 500 µL of human milk by ethanol precipitation. The proteins were then re-dissolved in 200 µL ammonium bicarbonate with 2% SDS and deglycosylated with 1000 units of PNGase F (New England Biolabs Inc., 1 unit defined as the amount required to remove >95% of the carbohydrate from 10µg of RNase B, denatured in 0.5% SDS, 40 mM DTT, 37 °C, pH 7.8). The high concentration of enzyme appeared to allow the use of SDS with success; however, such an excess of enzyme was not within the means of this project due to the high cost of the PNGase F and scaled up amounts of protein. To counteract the activity of SDS, non-ionic detergents such as Nonidet NP-40 or Triton X-100 when added in a 5-fold excess of SDS can be used to counteract the inhibitory activity of SDS though the exact mechanism is as yet unknown (Szabo et al., 2010).

Approach: To solubilise the proteins, the casein-depleted skim milk precipitate was first dissolved in 10 mM ammonium bicarbonate with 10% SDS (Section 3.2.3). The proteins were then reduced with a final concentration of 10 mM DTT incubated at 80 °C for 30 minutes followed by alkylation with a final concentration of 20 mM IAA at room temperature, in the dark for 45 minutes (Jensen et al., 2012). A five-fold volume equivalent of Triton X-100 was included before the addition of PNGase F (Roche, 10 units per mg, 1 unit defined as the

enzyme activity which hydrolyzes 1 nmol dabsyl fibrin glycopeptide within 1 min at 37 °C, pH 7.8).

Outcome: While proteins were dissolved and denatured, the high amount of detergents are known to interfere with many downstream separation processes such as liquid chromatography (LC) and online liquid chromatography-mass spectrometry (LC-MS). Therefore a method of removing the detergents (SDS + Triton X-100) was required.

3.2.5 Isolation of Released Glycans from Solution containing SDS and Triton X-100

The removal of detergent following denaturation and digestion is a complicated process and success is highly dependent on the detergents used. SDS and Triton X-100 are known to bind extremely well to proteins and have low enough molecular weight such that they cannot be simply removed by typical filtration or dialysis (le Maire et al., 2000). Hence, alternative methods were explored in order to remove SDS and Triton X-100 from the denatured milk proteins and the released *N*-linked glycan mixture (Section 3.2.4).

A previous study by Hager and Burgess (1980) reported the removal of SDS by acetone precipitation of protein samples. Complete precipitation of protein was found to occur between 3:1 and 5:1 acetone-to-buffer ratios and almost 100% of SDS was found in the acetone supernatant with less than 0.1% occurring in the protein pellet (Hager and Burgess, 1980). A study by Verostek *et al. (*2000) reported on the co-precipitation of oligosaccharides and proteins in solution using ice cold acetone. The addition of 4 volumes of ice-cold acetone was able to precipitate 97-99% of glycans into the protein pellet with 1-3% remaining in the supernatant and even very dilute glycans precipitating in 80% acetone. Verostek *et al.* (2000) also further reported that released glycans could be extracted from the protein- glycan pellet with 60% ice-cold aqueous methanol, in which glycans are soluble and proteins are not. Based on the above studies, acetone precipitation was used in this work to remove SDS-and Triton X-100 from the enzyme digestion solution while co-precipitating *N*-linked glycans and

proteins from the in-solution preparative scale release carried out in Section 3.2.4. *N*-linked glycans were then extracted from the pellet.

Approach: To the in-solution digestion performed in 250 μL volume, 1 mL of ice-cold acetone was added and the solution was left at 4 °C for at least 1 hour with maximum precipitation achieved when left at 4 °C overnight. The solution was centrifuged at 7000 g and 4 °C for 10 mins and the acetone supernatant was removed by pipette. *N*-linked glycans were extracted from the pellet by the addition of 500 μL of 60% ice-cold aqueous methanol and thoroughly mixing by pipette until the pellet was homogenized. The protein was pelleted again by centrifuging at 3500 g for 5 mins and the methanol supernatant was removed by pipette. The methanol extraction was repeated twice and extracts were combined, dried by vacuum centrifugation and analysed by PGC-LC-ESI-MS (Section 2.2.5) to establish the presence of released *N*-linked glycans.

Outcome: However, following the acetone precipitation and methanol extraction, *N*-linked glycans could not be detected by PGC-LC-ESI-MS (Section 2.2.5) analysis. Despite reports of complete detergent removal by both Hager and Burgess (1980) and Verostek *et al.* (2000), it was found that acetone precipitation and methanol extraction did not completely remove SDS and Triton X-100 from the released *N*-linked glycans in the PNGase F digestion mixture. The remaining presence of detergent resulted in complete loss of retention, and caused permanent damage, to the PGC column. Extensive washing of the column with high organic solvent eventually identified that SDS had bound to the column; however, column retention could not be restored with further washing. Since clearly, the removal of detergent was problematic, methods of *N*-linked glycan release without the use of detergents were then explored for the preparative recovery (Ruhaak et al., 2009; Verostek et al., 2000).

3.2.6 Enzymatic In-solution Release of *N*-linked Glycans without Detergent

Considering that the role of SDS was to dissolve and denature milk proteins, and the role of Triton X-100 was to counteract the effect of SDS, urea was identified as a possible alternative denaturing buffer. Urea is a powerful denaturant and solubilisier of proteins at concentrations above 2M that promotes unfolding by the interaction of urea with hydrophobic regions of proteins, disrupting intramolecular interactions in favour of hydrogen bonds with urea and water (Bennion and Daggett, 2003).

While not a specific inhibitor of PNGase F, urea is known to destroy the activity of enzymes at high concentrations. PNGase F is reported to be stable in 2.5M urea at 37 °C for 24 hours and retains up to 40% of its activity in 5M urea (Maley et al., 1989). However, it was found that neither 2.5M nor 5M urea were able to fully dissolve the casein-reduced skim milk protein precipitate (Section 3.2.3). When the concentration of urea was increased, the precipitate was able to dissolve in 8M urea; however, at such high concentrations PNGase F would be completely denatured.

Unlike detergents, urea does not bind specifically to proteins and can be removed relatively easily by size exclusion chromatography (SEC). After denaturing in 8M urea and reduction and alkylation with DTT and IAA, a buffer exchange was used to obtain the denatured and reduced proteins in enzyme friendly conditions. Buffer exchange was achieved using Zeba™ spin desalt columns (Thermo Scientific) which utilize the principle of gel permeation chromatography (GPC), a type of SEC. In GPC, separation occurs via porous beads packed into a column into which small molecules (e.g. urea) can enter and have increased retention while larger molecules (e.g. proteins) are excluded from the packing and elute quickly.

Approach: A stock solution of casein-reduced milk proteins was prepared by dissolving the dried and weighed pellet in 8M urea. For each release, a volume was removed and reduced and alkylated with DTT and IAA (Section 3.2.4). To perform the buffer exchange, the Zeba™

spin desalt columns (2mL) were washed three times by applying 2 mL of 10 mM ammonium bicarbonate and centrifuging at 1000 g for 2 minutes. The denatured proteins were then applied and the columns centrifuged at 1000 g for 2 minutes, eluting the proteins into 10 mM ammonium bicarbonate. In order to maximize enzyme efficiency, milk proteins were buffer exchanged with 10 mM ammonium bicarbonate pH 8.6, which is the optimum pH of PNGase F (Maley et al., 1989). The reaction volume was kept small at 250 μL to minimise travel time between molecules and increase the probability of enzyme-substrate interactions (Wyskovsky, 1998). The reaction mixture was gently shaken at 450 rpm to increase the reaction rate above that achieved by simple diffusive mixing (Jung et al., 2008). While PNGase F is active for up to 72 hours in optimum condition, the digestion was incubated at 37 °C for a minimum of 24 hours and a maximum of 36 hours to allow extensive deglycosylation without potential bacterial proliferation (Tarentino et al., 1985). Released *N*-linked glycans were isolated using acetone precipitation and methanol extraction described in Section 3.2.5. Methanol extracts were then dried by vacuum centrifugation and analysed by PGC-LC-ESI-MS (Section 2.2.5) to confirm the presence of released *N*-linked glycans.

Outcome: The successful release and isolation of *N*-linked glycans by the above approach was confirmed by PGC-LC-ESI-MS (Section 2.2.5). Analysis of the supernatant of the acetone precipitation of the digestion mixture by PGC-LC-ESI-MS (Section 2.2.5) found no trace of *N*-linked glycans indicating successful co-precipitation of released glycans from solution into the protein-glycan pellet. Acetone precipitation (Section 3.2.5) had the advantage of co-precipitation of released *N*-linked glycans and proteins into an easily manageable pellet. Ice-cold 60% aqueous methanol extraction (Section 3.2.5) recovered the released *N*-linked glycans from the pellet. The residual protein pellet would also allow for subsequent *O*-linked glycan release if required. Note that the 250 μL digestion volume allowed recovery of released *N*-linked glycans to be performed in a 1.5 mL Eppendorf tube, following the addition of 4 volumes of acetone.

It was discovered that the success of the acetone precipitation technique was affected by the amount of protein in the 250 μL reaction volume. When *N*-linked glycan release was attempted with 15 mg of protein in 250 μL, 4 volumes of acetone were unable to completely precipitate proteins. The result was a cloudy solution that did not form a pellet even after centrifugation at 15, 000 g, 4 °C for 30 minutes. Improved precipitation was eventually achieved for 15 mg of protein after the addition of 100 volumes of acetone (25 mL in total), kept at 4 °C overnight. However, this greatly increased the handling volume and while precipitation was eventually achieved, proteins could not be pelleted at centrifugation speed that did not cause damage to 50 mL Falcon tubes. To avoid the need for high speed centrifugation, and to simplify the procedure, precipitation of the milk protein and released *N*-linked glycans was best achieved in this work in a 1.5 mL Eppendorf tube with 5 − 10 mg of milk protein in a 250 μL reaction volume and 4 volumes of acetone for 1 hour at 4 °C. However, an additional high speed centrifugation step could be used in the procedure if larger volumes were required.

The inclusion of a buffer exchange in between protein solubilisation and denaturation in 8M urea and the subsequent addition of PNGase F allowed enzymatic glycan release without the need for detergent. The buffer exchange was able to capture urea, DTT, and IAA, and denatured proteins were eluted into 10mM ammonium bicarbonate while remaining completely dissolved. The 2mL Zeba™ spin desalt columns appeared to have an upper limit of ~20 mg of protein and optimal results were obtained with 10 mg of milk protein. Loading more than 20 mg of denatured protein caused clogging of the 2 mL columns and the eluted proteins had a turbid appearance indicating incomplete dissolution or proteins in colloidal suspension. However, larger columns are available from Thermo Scientific should the buffer exchange method be required for higher quantities of protein.
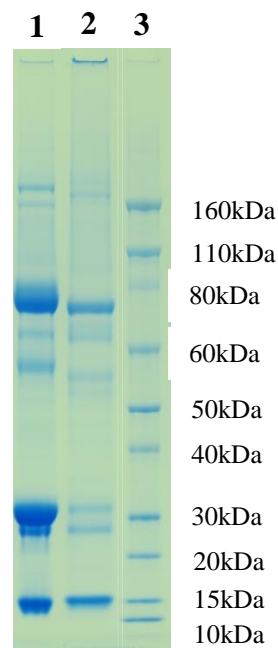
Figure 3.2 – Human milk protein precipitate before and after PNGase F deglycosylation separated on an SDS-PAGE gel stained with Coomassie Blue, lane 1: skim milk protein precipitate (~10 μg), lane 2: skim milk protein precipitate after de-glycosylation with PNGase F (~10 μg), lane 3: molecular weight markers.

De-glycosylation of proteins with PNGase F was monitored by SDS-PAGE (Figure 3.2). Successful release of *N*-linked glycans by PNGase F (Figure 3.2, Lane 2) was suggested for example by a reduction in heterogeneity and molecular weight of the ~80 kDa protein (Figure 3.2, Lane 1), which was most likely lactoferrin, an abundant human milk glycoprotein with three sites of N-linked glycosylation(Liu and Newburg, 2013; Wakabayashi et al., 2006). Since the heterogeneity of glycoprotein bands reflects the diverse range of glycoforms that are present, the reduction in the size and diffusion of bands can be explained by the loss of *N*-linked glycosylation. Furthermore, the position of casein bands remained unchanged which is consistent with the absence of *N*-linked glycosylation in caseins; however, the obvious decrease in heterogeneity in this range suggests the presence of other glycosylated proteins around this molecular weight such as the immunoglobulin chains (Froehlich et al., 2010). To validate the preparative scale release developed in this work, the released *N*-linked glycans were analysed by PGC-LC-ESI-MS (Section 3.2.7).

### 3.2.7 Comparison between the Analytical and Preparative Scale Releases

The analytical scale "dot blot" release (Section 2.2.4) of *N*-linked glycans from casein-reduced skim milk proteins was compared with the "in-solution" preparative scale release developed in this chapter.

Approach: The *N*-linked glycans released by both methods were analysed by PGC-LC-ESI-MS (Section 2.2.5) and the average mass spectra were compared (Figure 3.3).

Analytical scale: N-linked glycans were released from approximately 150 μg of human milk proteins by the "dot blot" method described in Section 2.2.4. Released *N*-linked glycans were dried by vacuum centrifuge, re-dissolved in 10 μL of MilliQ water and 2 μL were loaded for PGC-LC-ESI-MS.

Preparative scale: N-linked glycans were released from 10 mg of human milk proteins (obtained by method described in Section 3.2.3) using the method developed in Section 3.2.6. Released *N*-linked glycans were dried by vacuum centrifugation, re-dissolved in 100 μL of MilliQ water. An aliquot of 1 μL was removed and diluted with 9 μL of MilliQ water and 2 μL were loaded for PGC-LC-ESI-MS.

Outcome: The "in-solution" preparative scale release was validated by PGC-LC-ESI-MS (Figure 3.3). The average mass spectrum is the sum of all masses detected within the selected elution time frame and is able to provide an overview of the complete glycoprofile. The average mass spectra of the *N*-linked glycans released using both methods were almost identical and contained the same abundant masses indicating successful *N*-linked glycan release with little selective loss during glycan recovery. The only major difference was in intensity which was expected since approximately 100 fold more starting material was used in the preparative scale release.
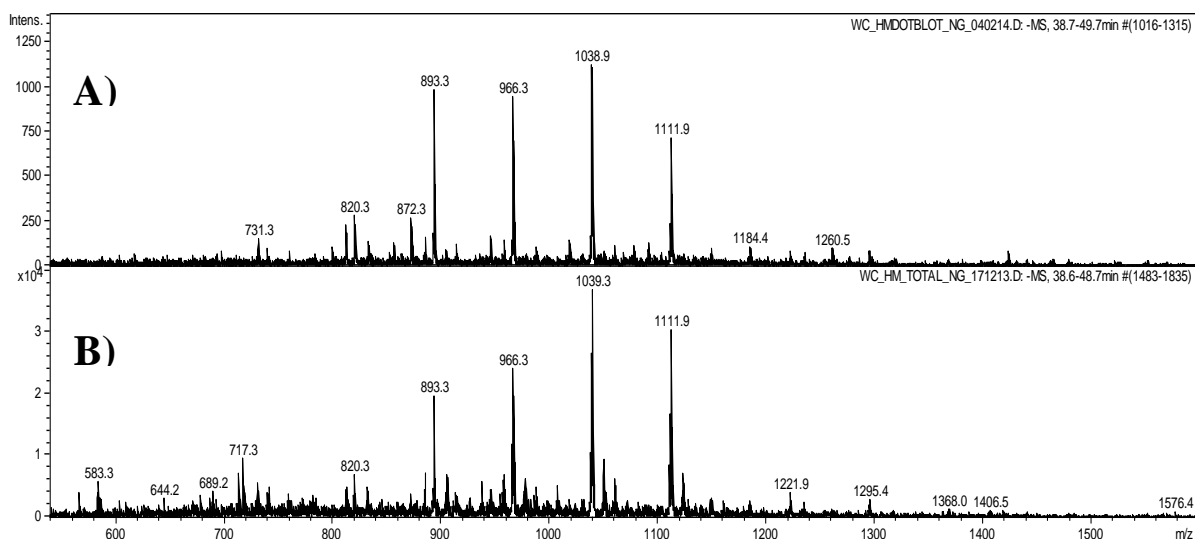
Figure 3.3 – Comparison of the average mass spectra of *N*-linked glycans released from human skim milk proteins using the analytical "dot blot" (Section 2.2.4) and preparative scale "in-solution" release. Panel A: *N*-linked glycans released by "dot blot" from 150 μg of milk protein; Panel B: *N*-linked glycans released by "in-solution" on preparative scale from 10 mg of milk protein diluted 1 in 1000. Average mass spectra obtained by PGC-LC-ESI-MS (Section 2.2.5).

While the "in-solution" preparative scale release was developed using human milk as a source of glycoproteins and *N*-linked glycans, the method can easily be applied to other biological samples or large quantities of purified glycoproteins with little optimisation required. In particular, glycoproteins which are soluble without using 8M urea could be denatured with DTT and IAA, allowing the buffer exchange (Section 3.2.6) to be omitted. The method developed is also compatible with *O*-linked glycan analysis because the proteins de-glycosylated with PNGase F are precipitated with acetone, and *N*-linked glycans are extracted from the pellet, leaving the *O*-linked glycans on the remaining protein pellet.

3.2.8 Quantitation of Released *N*-linked Glycans

Methods for the quantitation of large amounts of glycans are mainly limited to monosaccharide compositional analysis and colourimetric methods that are destructive and often lack sensitivity (Orlando, 2010; Varki, 2009). Monosaccharide compositional analysis involves the hydrolysis of glycans into constituent monosaccharides, which are then detected and quantified by colorimetric, gas chromatography (GC), HPAEC-PAD, or MS analysis (Varki, 2009). All these techniques are destructive and insufficient material was available for

these approaches to be used. Colorimetric stains such as the phenol-sulfuric acid and the periodic acid – Schiff's reagent (PAS) stain are commonly used to detect the presence of glycoproteins. The PAS stain converts vicinal diols of terminal monosaccharides into aldehydes that can then be reacted with Schiff's reagent for colour formation and measurement by UV absorbance (Leblond et al., 1957). However, Kilcoyne *et al.* (2011) developed a scaled down plate based assay using the PAS stain reaction, which provided a method for approximating the molar amount of glycans released from the casein-depleted skim milk proteins (Section 2.2.3).

Approach: The molar amount of human skim milk protein glycosylation was approximated with the PAS stain based plate assay developed by Kilcoyne *et al.* (2011) (Section 2.2.3). Bovine fetuin was used to generate a standard curve of absorbance at 550 nm (Figure 3.4) relative to the molar amount of *N*-linked glycosylation calculated from the number of moles of fetuin (~48.1 kDa) multiplied by three for the number of *N*-linked glycosylation sites.

$y = 165.82x + 159.81$
$R^2 = 0.9822$
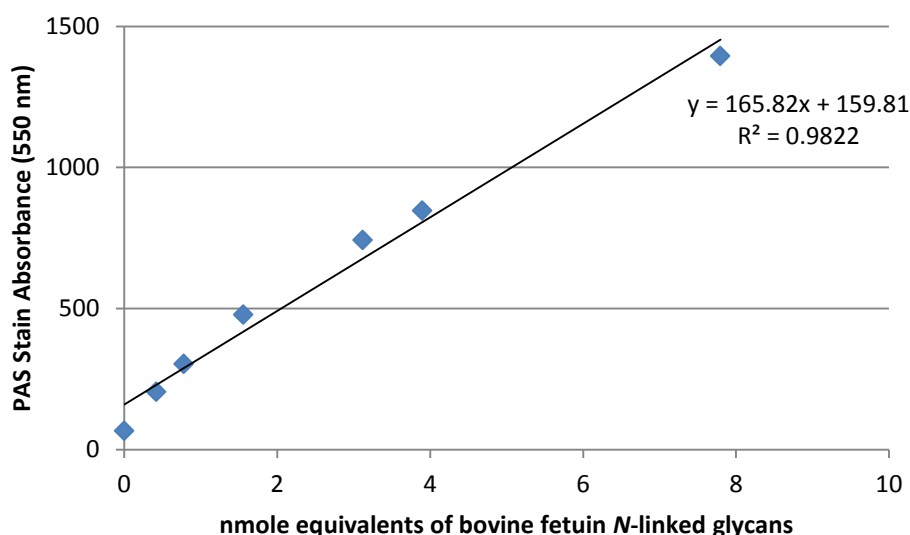
Figure 3.4 – Standard curve of the absorbance at 550 nm exhibited by PAS stained bovine fetuin

The PAS stain absorbance of 25 and 50 μg of casein-depleted human milk proteins (Section 3.2.3) was then compared against the fetuin standard curve to obtain an average of the molar amount of total glycosylation. For approximating the amount of *N*-linked glycans released by PNGase F, the assay was performed before and after deglycosylation with PNGase F (Section 3.2.6) to account for the PAS stain absorbance of *O*-linked glycans.

Outcome: Using the assay developed by Kilcoyne *et al.* (2011), the approximate molar amount of *N*-linked glycosylation on the casein-depleted human milk proteins was found to be ~125 nmol of *N*-linked glycans per mg of protein based on PAS stain absorbance when compared against the standard curve of bovine fetuin (Figure 3.4). Following the release of *N*-linked glycans with PNGase F, the molar amount of glycosylation remaining on the skim milk proteins was estimated to be ~40 nmol of presumably *O*-linked glycans per mg of protein. Subtraction of the amount of sugar on the total protein minus the sugar on the deglycosylated protein suggested that 85 nmol of *N*-linked glycans was released by PNGase F per mg of casein depleted human milk protein. Given that the milk used in this work contained an average of 2 mg/mL of non-casein protein, the amount of *N*-linked glycans per mL of milk was estimated to be ~170 nmol. However, this value only provides a very rough estimation as to the amount of *N*-linked glycans released and the absolute value of *N*-linked glycans obtained was not able to be determined due to the destructive nature of the quantitative methods available. Additional quantitation techniques that could be used are discussed further in Chapter 7.

3.3 Chapter Summary

A method of releasing *N*-linked glycans on a preparative scale was developed by scale up of the analytical scale dot blot to an in-solution release. Human skim milk was prepared from whole human milk by centrifugation at 10 000 g 4°C for 30 minutes. Insoluble caseins were precipitated by pH adjustment to 4.6 and the addition of 100 mM $CaCl_2$ and pelleted by ultra-centrifugation at 100 000 g, 4 °C for 90 minutes. The remaining skim milk proteins in the supernatant were isolated by chloroform/methanol precipitation (Section 3.2.2, Wessel and Flugge, 1984). Skim milk proteins were dissolved and reduced in 8M urea with 10 mM DTT then alkylated with 20 mM IAA. Denatured skim milk proteins were then buffer exchanged into 10 mM ammonium bicarbonate pH 7.8 before N-linked glycans were released using PNGase F. Released *N*-linked glycans were co-precipitated with proteins in ice-cold acetone and collected by centrifugation. The *N*-linked glycans were recovered by triturating the pellet with 500 μL of 60% ice-cold aqueous methanol and collecting the supernatant. A detailed description of this procedure is provided as a laboratory protocol in Appendix I.

A major aim of this project was to develop a protocol to obtain purified *N*-linked glycans to be used as analytical standards using human milk as a source and was approached as a two-part challenge. Part one was completed by the development of a preparative scale release from human skim milk proteins. Part two required the separation and collection of unique *N*-linked glycan structures and is described in Chapter 4.

Chapter 4 – Preparative Scale Separation and Detection of *N*-linked Glycans

4.1 Introduction

In order to obtain purified standards from the released *N*-linked glycans from human milk (Chapter 3), a method of separating *N*-linked glycans on a preparative scale was required. Chromatography is a technique commonly used in the analysis and purification of complex mixtures of compounds and has the potential for scale up from analytical to preparative scale separations. The aim of the work in this chapter was thus to develop a technique suitable for the separation of preparative scale quantities of *N*-linked glycans, based on an established technique regularly used for analytical scale separation and analysis of glycans by PGC-LC-ESI-MS (Section 2.2.5). Since glycans do not have a specific absorption in the ultraviolet region, and therefore cannot be detected by on-line UV detectors, their offline detection using matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) is described.

4.1.1 Liquid Chromatographic Separation of Glycans

Chromatography is a common analytical technique that permits the separation of complex mixtures of molecules. In liquid chromatography (LC), a liquid mobile phase, containing sample components to be separated, migrates through a stationary phase composed of solid particles packed into a column (Lough and Wainer, 1995). Separation by chromatography is achieved by differential distribution during migration and is dependent on the relative affinity of sample components for either the stationary or liquid mobile phase (Lough and Wainer, 1995). The development of high performance liquid chromatography (HPLC, formerly referred to as high pressure liquid chromatography) involved the use of increased pump pressure to push the liquid mobile phase through tightly packed stationary phases with small particle sizes (Lough and Wainer, 1995). In contrast to traditional LC, which relies on gravity

flow to pull the mobile phase through the column, HPLC greatly improves chromatographic separation and resolution.

Porous graphitised carbon (PGC) is a stationary phase composed entirely of spherical carbon particles of approximately 75% porosity (Pereira, 2008). PGC exhibits strong affinity for highly polar molecules and has the capacity to distinguish between structurally related isomers, making PGC well suited to the separation of glycans (Melmer et al., 2011; Ruhaak et al., 2009). Glycans are considered highly polar due to the number of hydroxyl groups attached to carbon atoms and exist as enantiomers and a range of structural isomers comprised of the same monosaccharides in different sequence and/or linkage (Section 1.1.1). While the exact mechanisms are still unknown (Melmer et al., 2011), retention on PGC has been described as a polar retention effect (PREG – polar retention effect on graphite) whereby molecules of increasing polarity show increased retention on the graphite surface (Pereira, 2008). Retention is further influenced by planarity and the presence of certain functional groups that interact with the flat graphite surface (Pereira, 2008). Since PGC is composed entirely of carbon and is not silica-bonded like other stationary phases, typically destructive conditions such as extreme pH and high temperature can be used without causing damage (Pereira, 2008).

Detection and analysis of glycans separated by PGC-LC is typically carried out by direct coupling to ESI-MS. The coupling of chromatography to mass spectrometry is extremely useful in reducing ion suppression effects in the mass spectrometer whereby highly abundant analytes can mask the presence of less abundant analytes (Jensen et al., 2012; Melmer et al., 2011; Ruhaak et al., 2009). Work by Pabst *et al.* (2007) has also demonstrated that PGC-LC-ESI-MS can be used to generate retention time libraries useful in the analysis of isomeric structures. Further work established a retention time correlation system used to determine specific structural features of glycans separated on PGC-LC based on reference glycans (Pabst et al., 2007; Pabst et al., 2010; Stadlmann et al., 2008). Due to the success of analytical

scale separations using PGC and the availability of a larger column, scale-up to a preparative

scale separation of *N*-linked glycans from milk using PGC-HPLC was considered for this

work.

4.1.2 Preparative Scale Separations on Porous Graphitised Carbon

Scale-up and optimisation of a chromatographic separation is highly empirical and involves a

combination of theoretical considerations and extensive trial-and-error (Rathore and

Velayudhan, 2002). Rathore *et al.* (2002) discusses two key concepts that are extremely

useful when attempting scale-up of chromatographic separations. Firstly, the *modes of*

interaction encompass the physicochemical interactions by which separation is achieved.

Examples include ionic interactions for ion exchange chromatography or hydrophobic

interactions for reversed phase chromatography. Secondly, the *modes of operation* represent

the various ways in which separation can be achieved and is the main focus of the scale-up

and optimisation process. Examples include parameters such as solvent system, ion pairing

agents, column temperature, flow rate and the use of either isocratic or binary gradients. The

most practical approach to scale-up involves maintaining certain parameters, which enables

the independent evaluation of other parameters (Rathore and Velayudhan, 2002). An equation

for the scale-up of chromatographic separations, given below, is able to calculate the flow rate

of a preparative separation (F_{prep}) based on the flow rate of an analytical separation ($F_{analytical}$)

in order to maintain the elution pattern observed during analytical separation (Rathore and

Velayudhan, 2002).

$$F_{prep} = F_{analytical} \cdot \frac{D^2\,prep}{D^2\,analytical}.$$

$D^2\,prep$ is the inner diameter (i.d.) of the preparative column squared and $D^2\,analytical$ is the

i.d. of the analytical column squared. A condition of the equation is that column length and

particle size is kept constant, which when met, allows the same analytical solvent system and

gradient to be used in the preparative separation. Other constraints that should be considered

are column capacity and maximum pressure. No information on column capacity could be obtained from the manufacturer of the PGC columns used in this work. One study was found that estimated the loading capacity of a PGC column to be between 5 – 50 μg of solute per gram of stationary phase depending on the chemistry of the solute (Echols et al., 1998). PGC is a great advance on early carbon stationary phases, which were too fragile for routine LC applications and suffered from low retention capacity and mass transfer due to the irregularity of micropores, variety of functional groups, and presence of impurities (Pereira, 2008). The development of PGC in 1982 by Knox and Gilbert resulted in a method of producing uniformly mesoporous (containing pores with diameters between 2 and 50 nm) graphite by depositing carbon onto a silica gel template and heating the material to temperatures above 2000 °C. This process effectively removed the irregular micropores of the carbon, causing the material to become graphitised and possessing the mechanical stability and surface area required for HPLC separations (Gilbert and Knox, 1981). An early study of PGC packings for HPLC noted that such columns should possess particle strength able to withstand pressures up to 300 – 600 bar (Unger, 1983).

A preparative scale separation of *N*-linked glycans released from human milk proteins (Chapter 3) was attempted by PGC-HPLC with the aim of processing enough material for subsequent use as *N*-linked glycan standards to facilitate detailed structural characterisation and establish structurally specific retention time libraries. In addition, it was necessary to establish an effective off-line detection system for the glycan fractions separation by PGC-HPLC. The standard method of detection used for analytical scale separation and analysis of glycans, LC-ESI-MS, was not appropriate due to the destructive nature of detection. Although a split system from the LC into the mass spectrometer could have been constructed for simultaneous detection and collection, MALDI-TOF MS was explored as an alternative, rapid, off-line technique that could be used for glycan detection following fractionation of the *N*-linked human milk glycans on a preparative scale.

4.2 Optimisation of a Preparative Scale Separation of *N*-linked Glycans

Separation of *N*-linked glycans on a preparative scale by PGC-HPLC was attempted based on the established analytical scale separation technique of PGC-LC-ESI-MS (Section 2.2.5) as a frame of reference. A number of factors were considered including reducing-end state of the glycan, elution gradient, flow-rate, collection of fractions, and method of detection.

4.2.1 One-pot Deamination and Reduction of *N*-linked Glycans

Enzymatic glycan release from glycoproteins with PNGase F cleaves *N*-linked glycans from asparagine leaving an amine attached to the anomeric carbon at the oligosaccharide reducing-end (Figure 4.1 A). The product of this cleavage is considered a non-reduced glycosamine, which has the advantage of potential derivatisation to introduce a chromophore allowing online detection when coupled to HPLC systems. However, once derivatised a glycan is no longer a true representation of its native form and may have loss of modification during the derivatisation reaction, altered column retention, fragmentation when analysed by MS, and biological activity *in vitro* and therefore has limited use as a standard for underivatised glycan analysis and in biological assays. In addition, PGC is able to separate the reducing end α- and β-anomers of the same structure thus effectively doubling the number of LC peaks to no advantage (Alpert et al., 1994; Varki, 2009). Therefore, it is common practice to convert the reducing end monosaccharide to an alditol thereby preventing mutarotation between the α- and β-anomers (Jensen et al., 2012; Zaia, 2008).

A schematic for the reactions involved in the deamination and reduction of *N*-linked glycamines is shown below (Figure 4.1). The reducing end amine is first removed (Figure 4.1 B) before reduction with sodium borohydride under alkaline conditions to form vicinal diols (Figure 4.1 C).
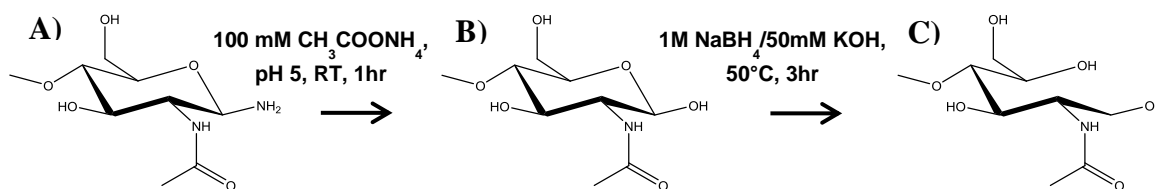
Figure 4.1 – Schematic of the established analytical scale technique for the reduction of *N*-linked glycans released by PNGase F cleavage (Section 2.2.4). (A) A released *N*-linked glycan carrying the Asn amine; (B) The amine attached to the anomeric carbon is removed with ammonium acetate under acidic conditions; (C) The anomeric carbon is then reduced to vicinal diols with sodium borohydride under alkaline conditions.

The *N*-linked glycans released from human milk proteins (Chapter 3) were therefore reduced prior to MS analysis and possible collection. However, the established protocol in the laboratory for *N*-linked glycan reduction (Section 2.2.4, Figure 4.1) has been optimised for the amount of *N*-linked glycans released from only microgram quantities of protein on a dot blot. To optimise on a larger scale, the reduction reaction was first optimised using the *N*-linked glycans from the standard glycoprotein bovine fetuin, which had been released using the preparative scale method developed in Chapter 3 (Section 3.2.6), before applying the method to the milk *N*-linked glycans.

Approach: A preparative scale reduction was first attempted using the conditions of the established analytical scale reduction method (Section 2.2.4, Figure 4.1), simply with increased volumes. The *N*-linked glycans released from 10 mg of bovine fetuin were dissolved in 100 µL of MilliQ water and deaminated with 20 µL of 100 mM ammonium acetate (prepared by the addition of acetic acid to MilliQ water and adjusting to pH 5 with ammonia), incubated at room temperature for 1 hour then dried by vacuum centrifugation. The *N*-linked glycans were reduced with 100 µL of 1 M sodium borohydride in 50 mM potassium hydroxide and incubated at 50 °C for 3 hours. The reaction was quenched by the addition of 10 µL of glacial acetic acid and the reduced *N*-linked glycans were purified by solid phase extraction (SPE) using graphitised carbon (HyperSep Hypercarb, Thermo Scientific), which is an established method of obtaining glycans from high salt containing solutions (Packer et al., 1998). The column was conditioned using three 2 mL volumes of

80% (v/v) acetonitrile and three 2 mL volumes of MilliQ water. The reduction reaction mixture was then applied to the column by gravity flow to maximise retention and the column was washed with 2 mL of MilliQ water three times to remove salts. The *N*-linked glycans were then eluted from the PGC SPE column with 40% (v/v) acetonitrile, 0.05% (v/v) TFA.
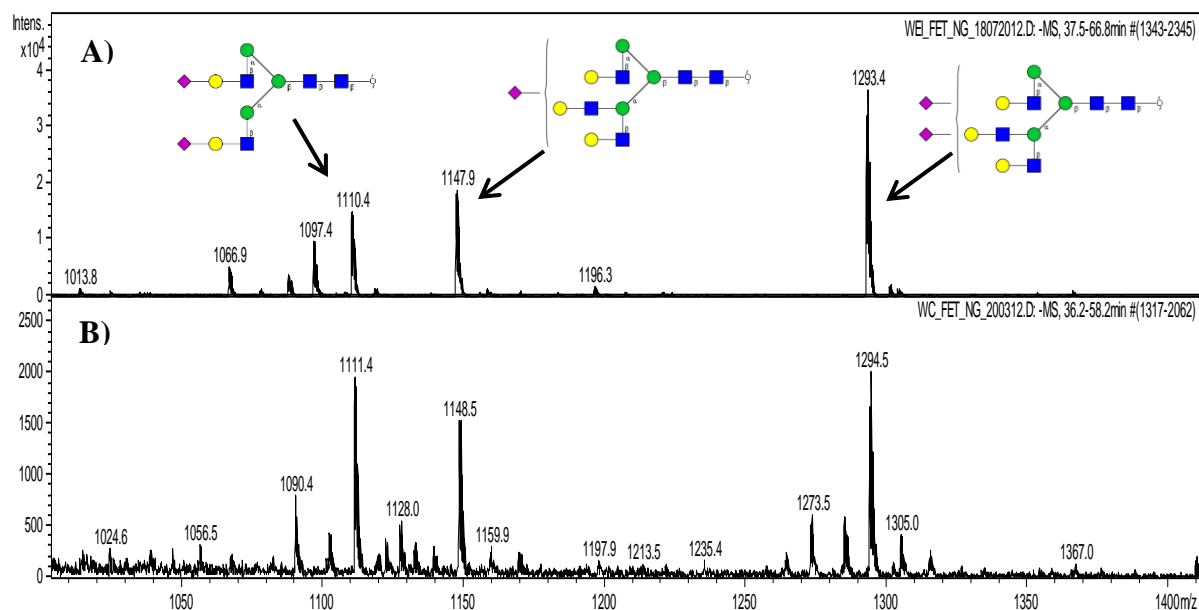
The reduced *N*-linked glycans from fetuin were analysed by PGC-LC-ESI-MS (Section 2.2.5) and compared against the starting material of unreduced *N*-linked glycans to check the completeness of the reaction. However, the procedure described above appeared to result in the complete loss of *N*-linked glycans since none could be detected by PGC-LC-ESI-MS analysis of the reduced sample. It was postulated that during the removal of ammonium acetate by vacuum centrifugation, a gradual decrease in pH may have taken place, which could have destroyed the glycans by acid hydrolysis; the larger volume of 120 µL used in the preparative scale reduction required at least one hour to completely dry by vacuum centrifugation compared to the analytical scale volume of 60 µL, which usually takes no longer than 15 mins. The glycans thus remained in low pH conditions for an extra 45 mins. Considering that sodium borohydride in alkaline conditions was required for a complete reduction reaction it was hypothesised that the vacuum centrifugation step to remove ammonium acetate could be omitted as long as an alkaline pH was attained. A "one-pot" reduction reaction was thus attempted, as described below.

"One pot" reduction: The *N*-linked glycans released from bovine fetuin were dissolved in 100 µL of MilliQ water and deaminated with 20 µL of 100 mM ammonium acetate pH 5 incubated at room temperature for 1 hour. The sample was not dried at this stage. The *N*-linked glycans were reduced in the same mixture by the addition of 120 µL of 2M sodium borohydride in 100 mM potassium hydroxide and incubated at 50 °C for 3 hours. The reaction was quenched by the addition of 10 µL of glacial acetic acid and desalted by PGC SPE as above (Packer et al. 1998).The *N*-linked glycans were eluted with 40% (v/v) acetonitrile,

0.05% (v/v) TFA from the PGC SPE column and dried. The completeness of the reduction reaction was monitored by PGC-LC-ESI-MS (Figure 4.2) and compared against unreduced *N*-linked glycans.

Outcome: Following the one-pot reduction of *N*-linked glycans according to the above procedure, purification by graphitised carbon SPE was complicated by excessive foaming of hydrogen released from the large quantity of sodium borohydride acidified and when applied to SPE columns often lead to overflow out of the top of the column. Air pockets in the packing bed were also formed and possibly may have resulted in sample loss. Nevertheless, SPE with graphitised carbon was still able to recover the reduced *N*-linked glycans with minimal salt contamination.

The PGC-LC-ESI-MS spectrum of the released *N*-linked glycans from bovine fetuin prior to reduction is shown in Figure 4.2 (panel A). Direct addition of sodium borohydride in the one-pot reduction appeared to reduce the bovine fetuin *N*-linked glycans, indicated by a 2 mass unit increase in mass corresponding to an *m/z* increase of 1 mass unit of the doubly charged species 1110.4 and *m/z* 1293.4 (Figure 4.2, panel A) to *m/z* 1111.4 and *m/z* 1294.5 respectively (Figure 4.2, panel B). The structure of *m/z* 1147.9 (Figure 4.2, panel A) is incorrectly peak peaked and actually has a monoisotopic mass of *m/z* 1147.4. The detection of *m/z* 1148.5 (Figure 4.2, panel B) corresponded to complete reduction of this glycan.

Figure 4.2 – PGC-LC-ESI-MS average mass spectra of N-linked glycans released from bovine fetuin on a preparative scale. A) Unreduced *N*-linked glycans from bovine fetuin; B) Reduced *N*-linked glycans from bovine fetuin reduced using the one-pot method. Masses shown are the doubly charged [M - 2H]⁻ species.

Due to the successful reduction of *N*-linked glycans from bovine fetuin, the one-pot reduction was considered a method suitable for the reduction of the *N*-linked glycans released from human milk proteins on a preparative scale (Section 3.2.6).

4.2.2 Detection of *N*-linked Glycans by MALDI-TOF MS

One of the main aims of the work described in this chapter was to develop a method of *N*-linked glycan separation suitable for preparative scale amounts (Section 4.2.3). However, it was first necessary to establish an effective method of detecting the *N*-linked glycans separated by PGC-HPLC on a preparative scale. Glycans are non UV absorbing and do not contain a fluorophore unless derivatised, therefore chromatographic detection of glycans on an analytical scale typically requires derivatisation with UV absorbing/fluorescent labels or online coupling to other methods such as ESI-MS (Harvey, 2000; Ruhaak et al., 2010). Derivatisation was not used in this work for the detection of glycans on a preparative scale as the aim was to attain reduced glycans for potential use as native glycan standards for MS quantitation and biological assays, and the reduction reaction prevented further derivatisation.

70

Analytical scale glycan separation and analysis by LC coupled to ESI-MS (LC-ESI-MS, eg. Figure 4.2) involves an LC step that is intrinsically required for separation and subsequent detection by electrospray ionisation (ESI) (Chapter 1.2.2) and is optimised for extremely small flow rates (~2 μL/min) due to the high sensitivity of modern mass spectrometers. However, following the off-line preparative scale LC separation of the *N*-linked glycans from human milk a rapid simple detection method was required to monitor the fractions collected.

Therefore, an off-line detection method was developed using matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS, Section 1.2.4) to rapidly detect the *N*-linked milk glycans in the collected fractions following PGC-HPLC. MALDI-TOF MS has been used extensively in glycan analysis and many established methods are available in the literature (Harvey, 2003; Papac et al., 1998; Stahl et al., 1994; Wang et al., 1999). MALDI-TOF analysis is highly empirical and user intensive as the direction of the laser is controlled by the operator on a purely visual basis. Due to the heterogeneous formation of crystals, "hot spots" are often formed which contain a greater concentration of analyte, a process which is highly dependent on the matrix used, matrix solution and sample preparation.

While a number of matrices are available, 2,5-dihydroxy benzoic acid (DHB) is the most commonly reported matrix and has consistently been used in the analysis of neutral *N*-linked glycans (Choi and Ha, 2006; Harvey et al., 2008; Stahl et al., 1994; Wang et al., 1999). Neutral glycans are most commonly analysed by MALDI-TOF MS in positive ion mode. The absence of a basic site on a glycan inhibits protonation in MALDI thus the major ions generated by *N*- and *O*-linked glycans are the singly charged metal ion adducts $[M + Na]^+$ and the $[M + K]^+$ adducts, producing 2 peaks for each mass, which can complicate analysis due to the overlap of masses. To overcome this issue, the addition of sodium ions in the form of sodium salts can suppress the formation of the potassium ion (Choi and Ha, 2006).

Acidic glycans are more difficult to detect by MALDI-TOF MS since NeuAc residues are extremely labile and are often lost during ionisation (Papac et al., 1996). Permethylation is often employed to stabilise the labile nature of NeuAc by methylation of carboxylic acid residues (Papac et al., 1996). However, like other forms of derivatisation, permethylation was avoided in this work. The use of the matrix 2',4',6'-trihydroxyacetophenone (THAP) has been reported in a number of studies as a suitable matrix for acidic glycans analysed in linear negative ion mode without the need for permethylation (Hao et al., 2010; Morelle et al., 2009; Papac et al., 1996) and thus was tested in this study for its suitability for the analysis of the acidic milk glycans

The mixture of *N*-linked glycans was separated into neutral and acidic fractions prior to detection by MALDI-TOF MS to allow the use of different matrices and ionisation modes for neutral and acidic glycans. Packer *et al.* (1998) demonstrated that glycans retained on graphitised carbon during SPE could be separated into neutral and acidic fractions by sequential elution with 20% (v/v) acetonitrile for neutral glycans, and 40% (v/v) acetonitrile with trifluroacetic acid (0.05%, v/v) to obtain the remaining acidic glycans. Alternatively, the latter solvent could be used to elute all glycans retained on the column. Sequential elution was used to obtain the neutral and acidic *N*-linked glycans released from human milk proteins following reduction by the one-pot method and purification by graphitised carbon SPE (Section 4.2.1). The neutral and acidic *N*-linked glycans were then used to develop a method of detection using MALDI-TOF MS. Once established, the method was to be used for off-line detection during preparative scale separation of the *N*-linked glycans released from human milk (Section 3.2.6) by PGC-HPLC (Section 4.2.3).

Approach: *MALDI-TOF sample preparation:* Neutral *N*-linked glycans were crystallised with DHB prepared by dissolving 10 µg/µL of matrix in 50 – 70% (v/v) aqueous ethanol with 10 mM sodium acetate. Acidic *N*-linked glycans were crystallised with THAP prepared by

dissolving 10 μg/μL of matrix in a 1:1 ratio of acetonitrile and 10 mM ammonium citrate (Papac et al., 1996). Samples were prepared on the target by the dried droplet technique; 1 μL of matrix was first applied to a spot on the MALDI target plate then 1 μL of sample was applied to the matrix and allowed to dry at ambient conditions.

MALDI-TOF instrument parameters: MALDI-TOF was performed on an Applied Biosystems 4800 *Plus* MALDI TOF/TOF Analyzer operating in reflectron-positive ion mode for neutral *N*-linked glycans and linear-negative ion mode for acidic *N*-linked glycans. Calibration of the instrument was performed according to standard operating procedures with a mixture of four peptides (Bradykinin, Neurotensin, Angiotensin I, ACTH) used as an internal standard for positive ion mode, and a tryptic digest of the protein ovalbumin from chicken egg for negative ion mode. Ions below *m/z* 800 were suppressed. Each mass spectrum was generated by summing the ions of 600 laser shots.

Maltoheptaose was used as a standard to confirm the production of the $[M + Na]^+$ ions in positive ion mode. In order to validate *N*-linked glycan detection by MALDI-TOF MS, the *N*-linked glycans released from human milk proteins (Section 3.2.6) were first analysed by PGC-LC-ESI-MS (Section 2.2.5) to obtain an average mass spectrum of the glycans present in the mixture. A list of masses corresponding to the $[M + Na]^+$ adducts of the most abundant masses detected by PGC-LC-ESI-MS was then calculated (Table 4.1). *N*-linked glycans released from human skim milk proteins were then analysed by MALDI-TOF MS and the abundant masses in the spectra were matched against the expected masses as shown by PGC-LC-ESI-MS analysis.

Outcomes: Despite reports of THAP being successfully used as a matrix in linear negative ion mode MALDI-TOF MS analysis of acidic *N*-linked glycans, in this work negative ion mode MALDI-TOF MS using the matrix THAP suffered from extremely low ionisation with major inconsistencies between spectra of the same sample.

The neutral *N*-linked glycans from human milk proteins were significantly easier to detect. Matrix solvent was found to be a major contributor to DHB crystal formation during analysis of the neutral *N*-linked glycans; concentrations of less than 50% (v/v) ethanol or acetonitrile produced larger crystals while concentrations greater than 50% (v/v) produced smaller and more uniform crystals. This was most likely due to high concentrations of organic solvent evaporating quickly resulting in smaller crystals. The smaller more uniform crystals obtained with 70% (v/v) aqueous ethanol were found to be the most suitable for MALDI-TOF analysis and generated spectra with low background noise and good resolution, as demonstrated by the monoisotopic distribution of maltoheptaose in which defined peaks with a difference of precisely 1 mass unit were observed (Figure 4.3, red trace). The addition of 10 mM sodium acetate to the DHB solvent resulted in complete suppression of the potassium adduct of maltoheptaose (Figure 4.3, *m/z* 1191.7, blue trace), with only the sodium adduct (*m/z* 1175.7) observed.
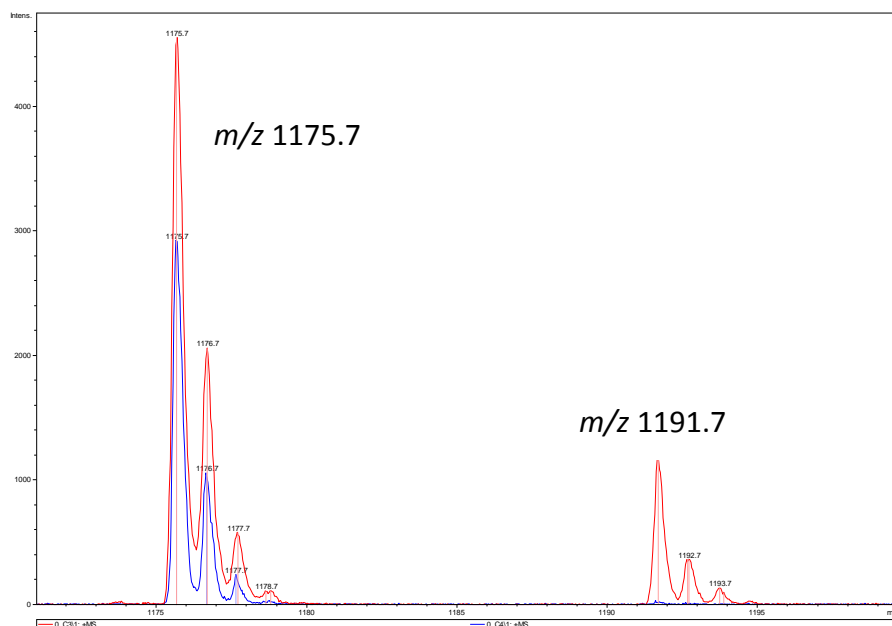

Figure 4.3 – Mass spectrum obtained by MALDI-TOF MS of maltoheptaose crystallised with 10 µg/µL DHB in 70% ethanol with (blue trace) and without (red trace) 10 mM sodium acetate. Good resolution was demonstrated by the monoisotopic distribution seen in the *m/z* 1175 peak.

The detection of neutral *N*-linked glycans by MALDI-TOF MS was validated by comparing the mass spectra of neutral *N*-linked glycans analysed by PGC-LC-ESI-MS and MALDI-TOF

MS (Figure 4.4). The most abundant masses from the ESI mass spectra that were detected as doubly charged ions in negative ion mode were used to predict the mass of the sodium adduct in the positive ion MALDI-TOF mass spectra (Table 4.1)

Table 4.1 – List of the abundant masses obtained from the ESI-MS average mass spectrum of neutral N-linked glycans released from human milk proteins with the charge state, native mass and predicted [M + Na]+ adduct mass.

m/z	Charge	Native mass	[M + Na]+
759.4	2	1520.8	1544.6
832.9	2	1667.6	1690.6
893.4	2	1788.6	1811.6
913.9	2	1829.8	1852.8
966.4	2	1934.8	1957.8
1039.5	2	2080.8	2103.8
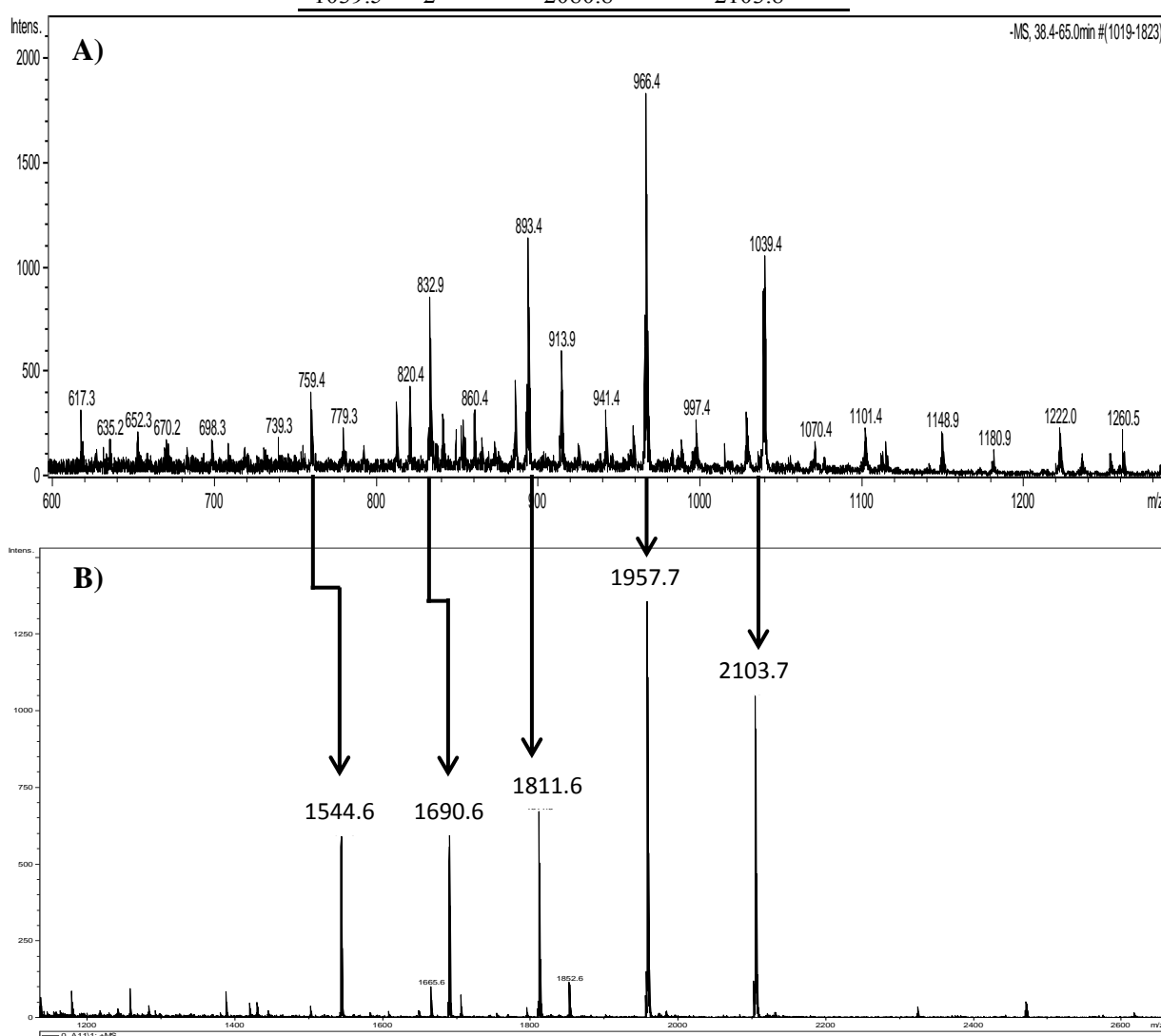
Figure 4.4 – The neutral *N*-linked glycans released from human milk proteins analysed by: (A) PGC-LC-ESI-MS detected as [M–2H]²⁻ doubly charged ions; (B) positive ion MALDI-TOF MS detected as [M+Na]⁺ singly charged ions. Other masses were also detected at low abundance (1665.6, 1852.6), indicating the presence of other glycans, the detection of which could have been suppressed by the more abundant masses.

The native mass (M) was calculated from the $[M - 2H]^{2-}$ doubly charged mass by multiplying the *m/z* by the charge state (x2) and compensating for the loss of 2 protons (+2 mass units). The mass of the $[M + Na]^+$ sodium adduct was then calculated by the addition of 22.9 mass units for a sodium ion.

The most abundant masses in the MALDI-TOF mass spectra (Figure 4.4, B) corresponded to the predicted $[M + Na]^+$ adducts (Table 4.1) accurate to within 1 Da, which confirmed the presence of neutral *N*-linked glycans released from human skim milk proteins and validated the use of the DHB matrix and positive ion mode MALDI-TOF MS as an offline detection method for neutral *N*-linked glycans.

4.2.3 Separation of Neutral *N*-linked Glycans by Porous Graphitised Carbon Liquid Chromatography

Having established an effective offline detection method for the neutral *N*-linked glycans (Section 4.2.2), it was then possible to proceed with the scaled-up separation of neutral *N*-linked glycans by liquid chromatography. Since an effective offline detection method for the acidic *N*-linked glycans was not able to be optimised, the neutral *N*-linked glycans were the focus of the preparative scale separation.

An analytical method of separating both *N*- and *O*-linked glycans by PGC-LC has previously been described by Jensen *et al.* (2012) using a 5 micron particle 0.18 mm i.d. x 100 mm size Hypercarb® column coupled to ESI-MS/MS analysis of glycans (Section 2.2.5). As this technique was already well established in the laboratory, it provided the basis for the scale-up using the method described by Rathore *et al.* (2002). With the availability of a larger Hypercarb® column in a 5 micron particle 2.1 mm i.d. x 100 mm size, the scale up equation (Section 4.1.2) was used to calculate a preparative scale flow rate of 272 µL/min, since the column length and chemistry of the stationary phase was maintained.

$$F_{prep} = 2 \; \mu L/min. \frac{(2.10)^2}{(0.18)^2}$$

$$F_{prep} = 272 \; \mu L/min$$

Approach: High performance liquid chromatography (HPLC) was performed on a Shimadzu LC-10AD VP Liquid Chromatograph unit with the solvent system and gradient described by Jensen *et al.* (2012) using 10 mM ammonium bicarbonate (Solvent A) and 45% acetonitrile in 10 mM ammonium bicarbonate (Solvent B). While slope of the gradient was maintained (Table 4.2) the program was altered slightly to include a longer period of 100% solvent A at the beginning and end of the separation. This was to allow at least 30 min for column equilibration for consecutive separations. When HPLC was performed at 272 µL per minute with water injection (10 µL MilliQ water), the column pressure was between 80 and 100 bar, which was well below the maximum recommended pressure of PGC columns (300- 600 bar, Section 4.1.1).

Table 4.2 – Gradient used for the preparative scale separation of neutral *N*-linked glycans on PGC.

Time	%B
0	0
15	0
15.1	5
60	35
80	100
85	100
86	0
100	0

Investigating column capacity: Increasing amounts of maltoheptaose were loaded and fractions were monitored for capacity of the column. Column overload is distinguished by sharp variations in retention time, elution in the void volume, peak broadening and peak splitting (Gritti and Guiochon, 2012). Upon loading 80 nmol (~80 µg) of maltoheptaose, column overload was observed by the elution of maltoheptaose within the void volume.

To determine the column capacity for the reduced *N*-linked glycans from milk (Section 4.2.1), a similar approach was taken by loading increasing amounts of the *N*-linked glycans and monitoring fractions. The column overload was observed upon loading the *N*-linked glycans released from ~1 mg of milk protein.

Fraction collection: Fractions separated by PGC-HPLC were collected by a Shimadzu FRC-10A automated fraction collector unit equipped with a custom built 96-well plate holder rack that was compatible with 250 μL well plates for a single separation, and 2 mL deep well plates for up to eight consecutive separations. By adjusting the flow rate of the preparative scale separation to 250 μL per min (a rate close to 272 μL per min as calculated by the scale-up equation), fractions spanning 1 minute could be collected in each well of the 96-well plate. For consecutive separations using 2 mL deep well plates, the 250 μL per min flow rate was maintained and the fraction collector was programmed to return to well A1 at 0 mins.

Glycan detection: All 96 fractions were removed from the plate and dried by vacuum centrifugation. Each fraction was dissolved in 50 μL of MilliQ water and an aliquot of 1 μL was removed for offline MALDI-TOF MS analysis (Section 4.2.2).

Outcome: Preparative scale amounts of *N*-linked glycans released from ~30 mg of human milk proteins were fractionated by PGC-HPLC, and detected by off-line by MALDI-TOF MS, equating to ~30 separations. Fractions from single separations were collected in 250 μL 96 well plates. Up to eight consecutive separations were possible using 2 mL deep well plates. Neutral *N*-linked glycans released from human skim milk proteins were detected in fractions 49 to 68 by MALDI-TOF MS (Section 4.2.2) with varying degrees of separation and purity (Figure 4.5). The mass spectra showed fractions 49, 57, 58, 60, 61, 62 and 63 contained a single peak indicating high purity (Figure 4.5, black squares) while the spectra from other fractions such as 55 and 64 contained several peaks indicating a mixture of glycans of different masses. In comparison with the MALDI-TOF mass spectra of the total neutral *N*-linked glycans released from human milk proteins (Figure 4.4, B), a greater number of masses

was detected after separation by PGC-LC (Figure 4.5) providing an excellent example of the ion suppression effect in the mixed sample, further reinforcing the practice of coupling chromatographic separation to mass spectrometry.

The elution pattern of the preparative scale separation (Figure 4.5) was compared against that of the analytical scale separation using the base peak chromatogram (BPC) obtained from PGC-LC-ESI-MS analysis of the same sample (Figure 4.6). The BPC represents the most intense signal at each time point providing the time frame and elution pattern of the analytical scale separation. As seen in Figure 4.6, the time frame of elution was 40 – 55 min as opposed to the preparative LC elution between 49 – 67 min. The two peaks observed in the BPC before 40 minutes (Figure 4.6, A and B) that exhibited high intensity, were identified as singly charged species that were too small to contain the chitobiose core and thus could not be considered *N*-linked glycans. Due to their small size, the two peaks could possibly be due to *O*-linked glycan contamination or as a result of ion adducts formation.

Glycan masses

[M+Na]+ m/z	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
1544	■	▓	░	░															
1706		▓	▓																
1389			░	▓															
1284				▓															
1690					■	▓													
1583						▓	░												
1745						▓	░												
1852						▓	▓												
1907						░	▓												
1957						▓	▓	■					■	■	░	░			
1811									■	░					▓	▓	░	░	░
1665										▓									
2103										░	■	■							
2322															░	░			
2470															░	░			
2101																	▓	▓	
2468															░				
2616															░				
1955																	░	░	▓
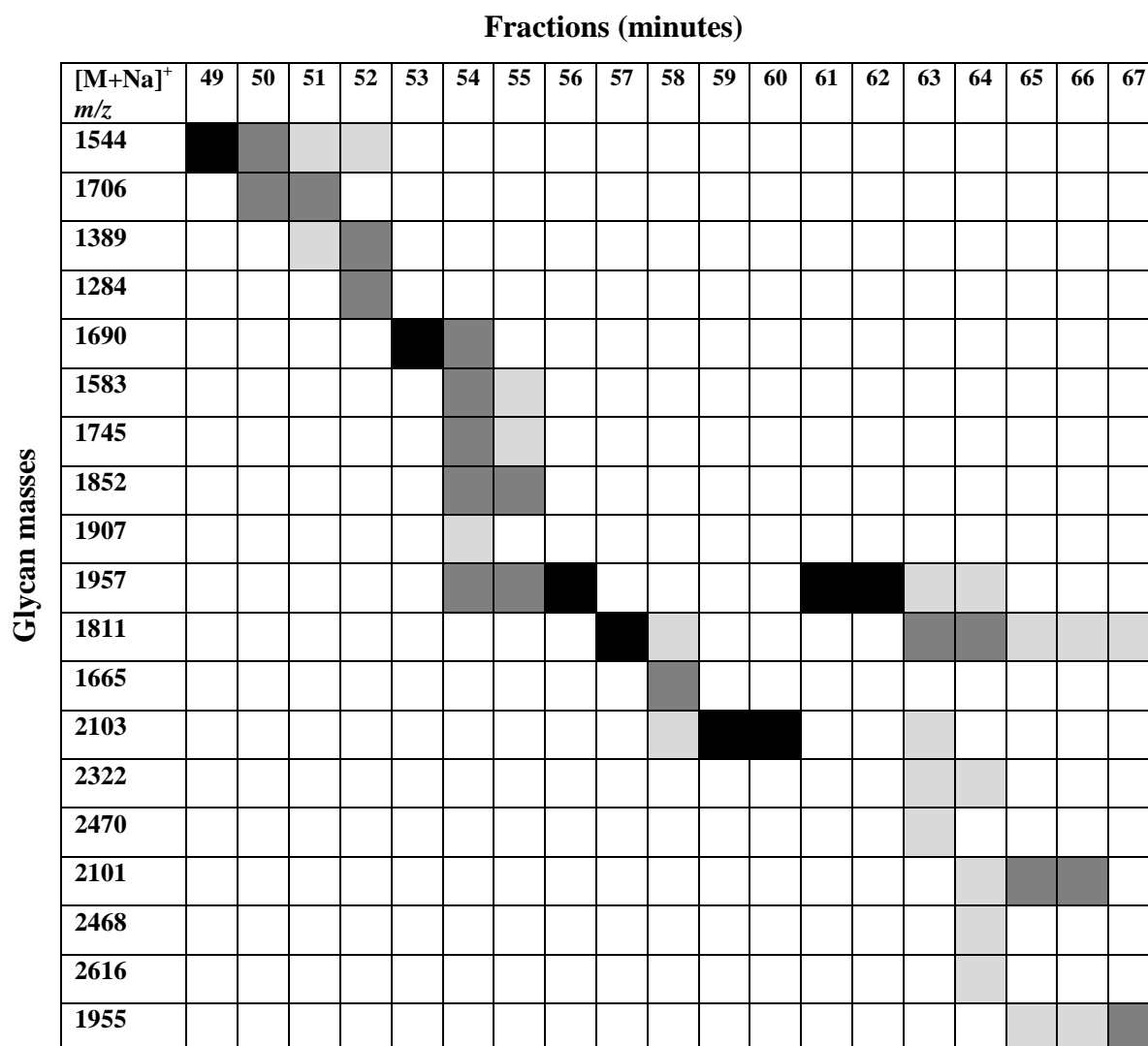
Figure 4.5 – Visual representation of the glycans separated by PGC-LC and detected by MALDI-TOF MS; fractions are presented across the top row, detected [M + Na]+ adduct masses are present down the left column. Black squares: collected fractions; Dark grey squares: present at high intensity; Light grey squares: present at low intensity. No glycans were detected in fractions eluting before 49 minutes or after 67 minutes.
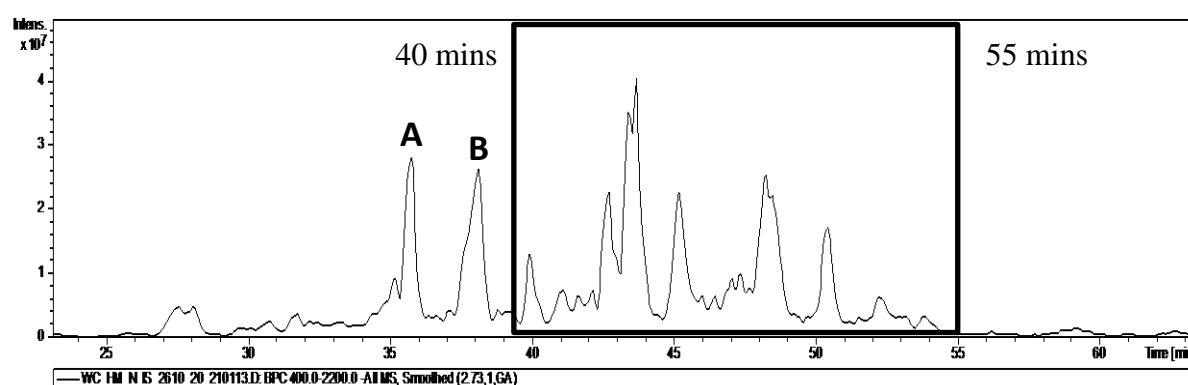
Figure 4.6 – Base peak chromatogram from LC-ESI-MS of the neutral N-linked glycans released from human milk proteins. The BPC provides the elution profile of the analytical scale separation. The majority of N-linked glycans eluted between 40 and 55 minutes (box).

During the preparative scale separation, the majority of *N*-linked glycans were detected eluting between 49 – 67 min (Figure 4.5). Scale up to a preparative scale separation thus resulted in a 7 min increase in the relative time frame of elution, with glycans starting to elute 9 mins later (Figure 4.5, 49 mins) in comparison to the analytical scale separation (Figure 4.6, 40 mins). A likely cause in of the shift in elution was the lower 250 μL per min flow rate rather than 272 μL per min as calculated using the scale up equation, which is designed to scale up separations while maintaining the elution pattern. By decreasing the flow rate, analytes in the mobile phase take longer to migrate through the column and take longer to elute. Another possible cause of this shift in retention time may have been the inherent differences between the Shimadzu HPLC system used in the preparative scale separation and the Agilent LC system used for analytical scale LC-ESI-MS method (Section 2.2.5). Factors such as the size and internal diameter of the tubing alone can produce significant differences in the void volume of the system, such that the solvent takes longer to travel from the mixer to the column resulting in a delayed elution profile. Nevertheless, despite the slight shift in the retention time, the relative time frame and pattern of elution was maintained.

The reproducibility of the preparative scale separation had a variation of ~1 min between separations performed at different times, which was seen in the MALDI-TOF MS spectra of collected fractions. A total of 15 individual separations were performed and similar fractions as determined by MALDI-TOF MS were combined. Reproducibility was greatly improved by equilibration of the column for at least 30 mins in 100% solvent A (10 mM ammonium bicarbonate) before sample injection. The extension of the 100% Solvent A segments at the beginning and end of the gradient allowed eight consecutive separations and collections to be performed automatically using 2 mL deep well plates. Two sets of eight consecutive separations were performed (16 separations in total) with no overlap of pure and mixed fractions.

The elution pattern appeared to be determined largely by size; larger masses generally eluted later than smaller masses (Figure 4.5). However, there were notable exceptions with instances of larger masses eluting earlier than smaller masses such as *m/z* 1544 and *m/z* 1706 eluting before *m/z* 1389 and *m/z* 1284 (Figure 4.5), and instances of the same masses eluting at different retention times (Figure 4.5, *m/z* 1957, *m/z* 1811). The differential elution of *m/z* 1975 and *m/z* 1811 in multiple fractions was most likely due to the separation of isomeric structures, which is an established capability of the PGC stationary phase (Pabst et al., 2007). The major outcome of these preparative scale separations was the purification of six *N*-linked glycans to >90% purity. In addition to providing a simple, rapid, off-line detection method, MALDI-TOF MS also allowed the monosaccharide composition of glycan masses to be determined as described below.

4.2.4 Compositional Analysis of Glycans Separated by Preparative Scale PGC-HPLC

A single stage of mass spectrometry (MS) using MALDI-TOF can be used to obtain the mass of a given glycan providing valuable information on its composition (Zaia, 2008). From mass measurement, the monosaccharide composition can be calculated by making informed assumptions based on the class of a glycan (*N*- or *O*-linked), possible monosaccharides (based on organism), and biosynthetic rules (Section 1.1.4) (Zaia, 2008). Compositional analysis is aided by bioinformatic tools such as GlycoMod, a program designed to calculate all possible compositions of a glycan from an experimentally determined mass (Cooper et al., 2001).

GlycoMod is available on the internet (http://web.expasy.org/glycomod/) within the ExPASy suite of proteomics tools and is linked to UniCarbKB (http://www.unicarbkb.org/), a knowledge base for glycomics and glycobiology research that contains experimentally determined glycan structures. When using GlycoMod, parameters such as *N*- or *O*-linked glycosylation, reducing end state (native, reduced or derivatised), and release with PNGase F or Endo (H or F), can be specified to refine the calculation.

Approach: Glycan compositions were calculated with GlycoMod using single stage MS data from the MALDI-TOF MS (Figure 4.5). GlycoMod search was performed with ± 0.5 Da mass tolerance and parameters were set to: Na^+ positive ion mode adduct, reduced *N*-linked oligosaccharides, and excluding monosaccharides not typically found in human secretions, i.e. pentose (xylose), sulfate, phosphate, *N*-glycolylneuraminic acid (NeuGc) and 2-keto-3-deoxynononic acid (KDN).

Outcome: GlycoMod was able to identify a number of masses as containing the (Man)3(GlcNAc)2 *N*-linked glycan chitobiose core (Table 4.3). Surprisingly, the masses *m/z* 1957 and *m/z* 2322 included NeuAc in their composition, which is the defining feature of acidic glycans and was not expected to be found in the neutral *N*-linked glycan fraction. In the case of *m/z* 1957, the potential NeuAc (291 Da) residue may also be indicative of two Deoxyhexose residues (292 Da), which differ in mass by one Dalton. The masses of *m/z* 1389, *m/z* 2470, *m/z* 2101, *m/z* 2616, and *m/z* 1955 could not be matched against any known *N*-linked glycan compositions. An expansion of the GlycoMod search parameters to include NeuGc suggested paucimannose compositions for *m/z* 2101, *m/z* 1955. Further expansion of the search parameters to include *O*-linked glycans resulted in a number of compositions for all five masses suggesting *O*-linked glycan contamination.

Table 4.3 – Masses of *N*-linked glycans detected by positive ion mode MALDI-TOF in order of elution with glycan compositions calculated by GlycoMod. Fractions collected at >90% purity are shown in bold. (Man)$_3$(GlcNAc)$_2$ represents the *N*-linked glycan chitobiose core. Hex = Glucose; Galactose; Man = Mannose; HexNAc = *N*-Acetyl-glucosamine, *N*-Acetyl-galactosamine; Deoxyhexose = fucose; NeuAc = *N*-acetylneuraminic.

MALDI-TOF $[M + Na]^{1+}$ *m/z*	GlycoMod Compositions	*N*-linked glycan class
1544	**(HexNAc)3 + (Man)3(GlcNAc)2**	**Complex**
1706	(Hex)1 (HexNAc)3 + (Man)3(GlcNAc)2	Complex
1389	No *N*-linked glycan match	
1284	(HexNAc)1 (Deoxyhexose)1 + (Man)3(GlcNAc)2	Complex
1690	**(HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2**	**Complex**
1583	(Hex)4 + (Man)3(GlcNAc)2	High Mannose
1745	(Hex)5 + (Man)3(GlcNAc)2	High Mannose
1852	(Hex)1 (HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2	Complex
1907	(Hex)6 + (Man)3(GlcNAc)2	High Mannose
1957	**(Hex)2 (HexNAc)2 (Deoxyhexose)2 + (Man)3(GlcNAc)2**	**Complex**
	(Hex)2 (HexNAc)2 (NeuAc)1 + (Man)3(GlcNAc)2	**Complex**
1811	**(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2**	**Complex**
1665	(Hex)2 (HexNAc)2 + (Man)3(GlcNAc)2	Complex
2103	**(Hex)$_2$ (HexNAc)2 (Deoxyhexose)3 + (Man)3(GlcNAc)2**	**Complex**
2322	(Hex)3 (HexNAc)3 (NeuAc)1 + (Man)3(GlcNAc)2	Complex
	(Hex)1 (HexNAc)1 (Deoxyhexose)5 (NeuAc)1 + (Man)3(GlcNAc)2	Complex
2470	No *N*-linked glycan match	
2101	No *N*-linked glycan match	
2468	(Hex)3 (HexNAc)3 (Deoxyhexose)3 + (Man)3(GlcNAc)2	Complex
	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2	Complex
2616	No *N*-linked glycan match	
1955	No *N*-linked glycan match	

4.3 Chapter Summary

As part of the aim to acquire purified *N*-linked glycan standards from human milk proteins, a preparative scale separation using PGC-HPLC with offline detection by MALDI-TOF MS was developed. In the optimised technique, the *N*-linked glycans released from human milk proteins (Chapter 3) were first reduced to glycan alditols using a one-pot de-amination and reduction reaction with 100 mM ammonium acetate and 2M sodium borohydride in 100 mM potassium hydroxide, and incubating at 50 °C for 3 hours (Section 4.2.1). The final protocol for reduction with sodium borohydride and preparative scale separation of neutral *N*-linked glycans by PGC-HPLC is provided in Appendix II. The reduced *N*-linked glycans were desalted by SPE on graphitised carbon and separated by sequential elution into neutral and acidic *N*-linked glycan classes with 20% (v/v) acetonitrile and 40% (v/v) acetonitrile with 0.05% (v/v) TFA respectively to allow separate analysis. MALDI-TOF MS was established as an effective offline detection method with the use of DHB matrix dissolved in 70% (v/v) ethanol with 10 mM sodium acetate, enabling identification of neutral *N*-linked glycans in fractions separated by PGC-LC on a preparative scale. Though MALDI spectra of the acidic *N*-linked glycans was unable to obtained in this project, a number of alternative approaches could be implemented including trial of different matrices, mixtures of matrices, solvents, and sample preparation. A total of 30 separations were performed using the *N*-linked glycans released from 30 mg of human milk proteins resulting in the collection of six unique *N*-linked glycan fractions of >90% purity. In order to use the six purified *N*-linked glycans as standards to establish an LC retention time library and facilitate structural analysis by MS, their structural characterisation by tandem mass spectrometry was continued in the work described in Chapter 5.

Chapter 5 – Structural Characterisation of *N*-linked Glycans from Human Milk for Potential Use as Glycan Standards

5.1 Introduction

The methods developed for preparative scale release (Chapter 3) and separation (Chapter 4) of the *N*-linked glycans from human milk enabled the isolation of six unique *N*-linked glycans of high purity for potential use as glycan standards. Possessing a standard of known structure permits the establishment of MS fragmentation and LC retention time libraries, which are extremely useful and lacking largely in the glycomics field in the characterisation of complex mixtures of unknown glycans, as well as the access to specific glycan structures for biological assays. In this chapter, detailed structural characterisation by PGC-LC-ESI-MS was performed to confirm the composition of the six purified *N*-linked glycan (Section 2.2.5). Additionally, a PGC-HPLC retention time library was established to facilitate their use as standards is described.

5.1.1 Structural Analysis of Glycans by Tandem MS Fragmentation

Whereas the mass gives an idea of the monosaccharide composition, a single mass often reflects a mixture of potential isomers and single stage MS analysis (such as MALDI-TOF MS) does not provide information on anomericity (α or β), linkage (C1 to C2, C3, C4 or C6) and branching configuration or epimeric forms of monosacharrides (Glu, Gal, Man). Therefore, Tandem MS (MS/MS) is necessary for the detailed structural characterisation of glycans (Leymarie and Zaia, 2012; Zaia, 2008). The principle of MS/MS involves the fragmentation of precursor mass ions. Fragmentation is usually achieved by collision-induced dissociation (CID), a process in which molecular bonds are cleaved by collision with gas atoms, which increases the vibrational energy of precursor ions to the point of rupture (Zaia,

2008). Bond rupture results in two types of cleavage: between sugar rings (glycosidic cleavage), or across rings (cross-ring cleavage). The mass spectra of resulting product ions enable structural characterisation by matching against previously described diagnostic ions (Everest-Dass et al., 2013; Harvey, 2005; Harvey et al., 2008; Zaia, 2010).

A systematic nomenclature for the fragmentation of glycans by tandem MS was introduced by Domon and Costello (1988, Figure 5.1). Product ions generated by glycosidic cleavage provide information on glycan composition and sequence, with ions containing the reducing end labelled as Y and Z, and ions containing the non-reducing terminus labelled as B and C (Domon and Costello, 1988). Ions are further distinguished depending on the location of the glycosidic oxygen following cleavage, which is present in C and Y ions and absent in B and Z ions (Figure 1.5). Cross-ring cleavage provides information on composition, branching and linkage and generates ions labelled as A and X and with the positions of ring cleavage provided in superscript. The location of glycosidic or cross-ring cleavage along the glycan chain is denoted in subscript beginning from the non-reducing terminus. The reducing end is noted by subscript R.
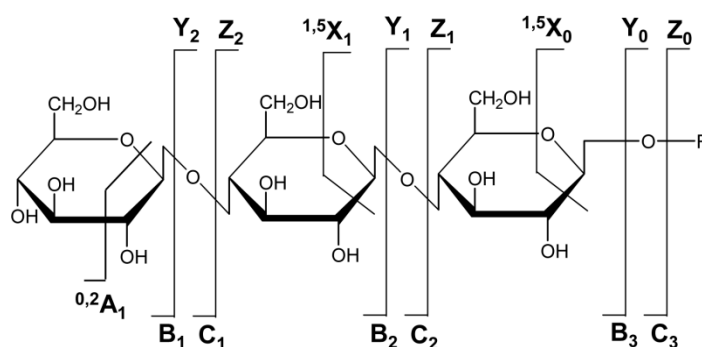
Figure 5.1 – Nomenclature used for the product ions of MS/MS fragmentation of glycans. The position of cleavage is given in subscript beginning from the non-reducing terminus. Position of cross-ring cleavage is noted in superscript. Nomenclature from Domon and Costello (1988), image from Zaia J. (2008).

Fragmentation is heavily influenced by the mode of ionisation (Harvey, 2005). Ionisation in positive ion mode produces mainly $[M + Na]^+$ and $[M + H]^+$ ions which fragment

87

predominantly by glycosidic cleavage. While this type of cleavage provides information on composition and sequence, positive ion spectra are considered more ambiguous due to the lack of cross-ring cleavage and information on linkage position. Alternatively, ionisation in negative ion mode produces $[M - H]^-$ ions, which fragment by both glycosidic and cross-ring cleavage (Harvey et al., 2008). Harvey *et al.* (2008) proposed that following the initial loss of a proton, the $[M - H]^-$ ion undergoes a series of electron shifts resulting in highly stable fragment ions that are diagnostic of specific structural features.

Harvey et al. (2008) described a number of prominent diagnostic ions useful in the structural characterisation of *N*-linked glycans by negative ion ESI-MS/MS. The D, E and F ions (Figure 5.2) provide structure specific information on the composition and linkage of the branched arms of *N*-linked glycans. In particular, the D type ion is a major diagnostic ion that provides specific information on the 6-arm of an *N*-linked glycan. The D ion is formed by the glycosidic cleavage of the core GlcNAc residues and the 3-arm, leaving the 6-arm plus a mannose and water molecule (+179 mass units) via C type cleavage (Figure 5.2). The presence of a bisecting GlcNAc can also be identified by an intense D − 221 ion resulting from the glycosidic cleavage of the β1-4GlcNAc from the D ion.
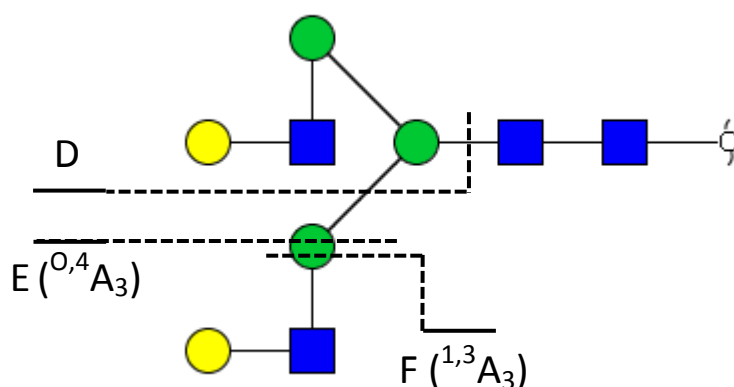
Figure 5.2 – Fragmentation of a bi-antennary *N*-linked glycan showing the cleavage of the D, E, and F diagnostic ions, adapated from Harvey *et al.* (2008). Monosaccharides are presented in the CFG format, Yellow circle: galactose; Green circle: mannose; Blue square: *N*-acetylglucosamine (GlcNAc).

The E type ion is described to be diagnostic of the 3-arm and is produced by ^{O,4}A cross-ring cleavage of the 3-arm mannose resulting in an ion containing the 3-arm plus carbons 101 mass units from the mannose residue (Figure 5.2, E). The E ion, in conjunction with the D ion, is most commonly used to identify the branching position of tri- and tetra-antennary glycans as branching of the 3-arm will produce an E ion containing both branches (Harvey et al., 2008).

The type F ion provides additional information on the composition of the arms and is produced by the cross-ring cleavage of a branching mannose residue leaving the arm plus 59 mass units for the 2 carbon atoms from the mannose (Figure 5.2, F). In Figure 5.2, the F ion shown is formed via ^{1,3}A cross-ring cleavage of a 2 linked branching mannose. Substitution of the mannose at the 4 and 6 positions can also produce F ions via ^{2,3}A and ^{O,4}A cross-ring cleavage respectively (Harvey et al., 2008).

Manual interpretation of MS data is greatly enhanced using GlycoWorkbench, a software tool developed by the EUROCarbDB initiative to assist the manual interpretation of MS data (Ceroni et al., 2008). GlycoWorkbench incorporates the GlycanBuilder (Ceroni et al., 2007), a visual editor that allows the assembly of glycan structures in a user friendly format. Monosaccharides can be displayed in a number of popular formats and stereochemical information such as anomericity and linkage position can be represented visually. Glycan masses can be computed and displayed taking into account reducing end state (native, reduced, or derivatised) and ionisation (negative or positive ion mode) with the identity and quantities of ion adducts (H^+, Na^+, Li^+, K^+, H^-, Cl^-, $H_2PO_4^-$).

GlycoWorkbench also incorporates both automated and manual *in silico* fragmentation of proposed structures that is capable of computing all topologically possible fragmentations based on the maximum number of glycosidic and cross-ring cleavages specified. The *in silico* fragmentation tool is particularly useful in the assignment of larger fragments ions by

allowing the user to cleave at specific positions on the proposed structure. The mass of the resulting fragments is calculated in terms of *m/z* and native mass aiding in the identification of singly or doubly charged ions. Software tools such as GlyoMod and GlycoWorkbench, greatly accelerate the interpretation of MS data and detailed structural characterisation is made possible by combining single stage and tandem MS/MS with established biosynthetic rules (Section 1.1.4), as defined by the documented activity of glycosyltransferases. However, even with CID fragmentation, specific information such as anomericity and the linkage position of terminal residues can still remain ambiguous. In such cases, sequential enzymatic digestion experiments with exoglycosidases can be performed to confirm anomericity and linkage position of terminal residues (Royle et al., 2007). A range of exoglycosidases with defined substrate specificities are commercially available that are capable of cleaving single monosaccharides of particular anomeric linkage and position. For example, β(1-4)-Galactosidase will specifically cleave galactose β1–4 linked to a GlcNAc while leaving galactose with β1–2 linkage.
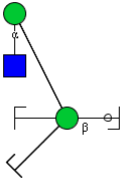
5.2 Methods
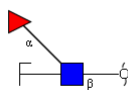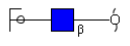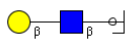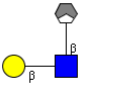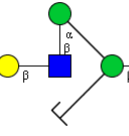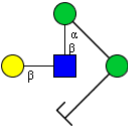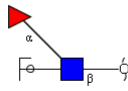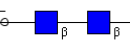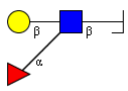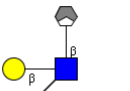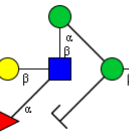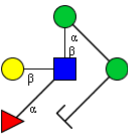
5.2.1 MS Data Analysis

N-linked glycans released from human milk (Chapter 3) and purified by PGC-HPLC (Chapter 4) were analysed in detail by tandem MS using PGC-LC-ESI-MS operating in the negative mode (Section 2.2.5). MS data was processed with ESI Compass 1.3 DataAnalysis Version 4.0 software. Glycan and fragment ions were present in the ESI MS spectra as both singly (denoted by 1- in superscript) and doubly (denoted by 2- in superscript) charged species.

5.2.2 Structural Characterisation of *N*-linked Glycans by MS/MS Fragmentation

Structural assignments were made using the MS/MS fragmentation spectra using diagnostic ions (Table 5.1) to identify key functional groups. Further interpretation of the spectra was facilitated by constructing the predicted glycan structure in GlycoWorkbench and performing *in silico* fragmentation. Theoretical fragments were then matched against the fragmentation spectra. The relative abundance of each composition or structure if determined was calculated by integration of the peak area of a mass in the extracted ion chromatogram (EIC). Relative abundance was presented as a percentage of the total of all compositions or structures when known.

Table 5.1 – Diagnostic ions used to identify specific functional groups. Ions are represented in the hybrid notation (Section 1.1.2) with the type of ion and *m/z* included (Everest-Dass et al., 2013; Harvey et al., 2008).

Core	Arms	D-ion

Z, *m/z* **350**

Y, *m/z* **222**

C, *m/z* **382**

F ^{1,3}A$_{3Man,}$ *m/z* **424**

D, *m/z* **688**

D, *m/z* **670**

Y, *m/z* **368**

Y, *m/z* **425**

B, *m/z* **510**

F ^{1,3}A$_{3Man,}$ *m/z* **570**

D, *m/z* **816**

D-18, *m./z* **834**

Bisecting GlcNAc

F ^{1,3}A$_{3Man,}$ *m/z* **262**

D-221, *m/z* **508**

5.2.3 Structural Confirmation by Exoglycosidase Digestion

The six *N*-linked glycans purified by PGC-HPLC (Chapter 4) were dried by vacuum centrifugation and taken up in 100 μL of MilliQ water. A mixture was prepared by combining 1 μL of each purified glycan and then dried by vacuum centrifugation. The following were added to the dried *N*-linked glycans: 2 μL of sodium acetate buffer pH 5 (5X), 2 μL of α(1-3,4)-Fucosidase from almond meal (Prozyme), 2 μL of β(1-4)-Galactosidase from *Streptococcus pneumonia*, and 4 μL of MilliQ water. The mixture was incubated at 37 °C overnight, purified by SPE on graphitised carbon (Section 2.2.4, *Desalting the reduced N-linked glycans*), then analysed by PGC-LS-ESI-MS (Section 2.2.5) and compared against an undigested mixture.

5.2.4 Establishment of a Retention time Library

The six *N*-linked glycan standards purified by PGC-HPLC (Chapter 4) were individually analysed by PGC-LC-ESI-MS (Section 2.2.5) to obtain retention times. In addition, a mixture was prepared by combining 1 µL aliquots of each purified glycan, which was then analysed by PGC-LC-ESI-MS (Section 2.2.5) to identify any shifts in retention time due to the presence of other glycans.

5.3 Results and Discussion

5.3.1 Detailed Structural Analysis by ESI-MS/MS

The compositions of the six purified *N*-linked glycans obtained using GlycoMod from the MALDI-TOF MS data in Chapter 4 (Section 4.2.4) were confirmed by PGC-LC-ESI-MS/MS and the detailed structures are shown in Figures 5.3 to 5.8. The glycan compositions of the singly charged $[M + Na]^+$ ions obtained by MALDI-TOF MS (Section 4.2.4) and corresponding doubly charged $[M - 2H]^{2-}$ ions obtained by PGC-LC-ESI-MS of the six purified *N*-linked glycans are provided below (Table 5.2). Two fractions of $[M + Na]^{1+}$ *m/z* 1957 were collected at different retention times from the PGC column and returned a GlycoMod composition containing either two Fuc or one NeuAc residues possibly suggesting structural isomers with different positions of substitution.

Table 5.2 – Masses of *N*-linked glycans detected by positive ion mode MALDI-TOF and negative ion mode ESI-MS in order of elution with glycan compositions calculated by GlycoMod. (Man)₃(GlcNAc)₂ represents the *N*-linked glycan chitobiose core. Hex = Glucose; Galactose; Man = Mannose; HexNAc = *N*-Acetyl-glucosamine, *N*-Acetyl-galactosamine; Deoxyhexose = fucose; NeuAc = *N*-acetylneuraminic.

MALDI-TOF $[M + Na]^{1+}$ *m/z*	ESI-MS $[M - 2H]^{2-}$ *m/z*	GlycoMod Compositions
1544	759.9	(HexNAc)3 + (Man)3(GlcNAc)2
1690	832.8	(HexNAc)3 (Deoxyhexose)1 + (Man)3(GlcNAc)2
1957 (2 isomers)	966.4	(Hex)2 (HexNAc)2 (Deoxyhexose)2 + (Man)3(GlcNAc)2
		(Hex)2 (HexNAc)2 (NeuAc)1 + (Man)3(GlcNAc)2
1811	893.3	(Hex)2 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2
2103	1039.3	(Hex)₂ (HexNAc)2 (Deoxyhexose)3 + (Man)3(GlcNAc)2

Major structural features such as the composition of the 3- and 6-arms were identified from the MS/MS fragmentation spectra by key diagnostic ions (Table 5.1). Core substitution was identified by the core related diagnostic ions (Table 5.1, Core). The presence of core fucosylation was indicated by ions corresponding to glycosidic cleavage of the reducing end GlcNAc linked to a fucose (Z_R, *m/z* 350^{1-}, Y_R, *m/z* 368^{1-}). The absence of core fucosylation

was indicated by ions corresponding to the loss of the reducing end GlcNAc (Y_R, m/z 222^{1-}) and the adjacent GlcNAc without fucose (Y_{R-1}, m/z 425^{1-}).

Though often low in abundance, the D ion was useful in directly identifying the composition of the 6-arm and indirectly identifying the composition of the 3-arm. The D ion was also often accompanied by the D-18 dehydration product indicating the loss of the glycosidic oxygen and 2 protons (Table 5.1, D ion). The presence of a bisecting GlcNAc was determined by a highly abundant D-221 diagnostic ion (m/z 508^{1-}, Table 5.1, Bisecting) which replaced a normal D and D-18 ion. The F type ions were also useful in elucidating the composition of the arms with the ^{1,3}A cleavage of the branching mannose producing abundant ions (Table 5.1, Arms).

Ions resulting from the ^{2,4}A cleavage of the GlcNAc adjacent to the reducing end were also present in high abundance and confirmed the composition of the non-reducing portion attached to the core mannose residues. Glycosidic cleavage of the terminal residues was also common producing both Y and Z ions and their corresponding B and C ions.

The m/z 759.9 and m/z 832.8 glycans were both identified as truncated (lacking galactose) bisected N-linked glycans (Figure 5.3 and 5.4) producing very similar fragmentation spectra differing only by ions diagnostic of core fucosylation. A prominent D-221 ion at m/z 508^{1-} ion was produced identifying the presence of a bisecting GlcNAc and a GlcNAc substituted on the 6-arm. Cross ring cleavage of the GlcNac adjacent to the reducing end produced an ^{2,4}A$_{R-1}$ ion that was present as m/z 576^{2-} and less abundant as m/z 1154^{1-}. Substitution of the core appeared to affect the abundance of these ions. In the spectra of the non-fucosylated glycan, the m/z 576^{2-} ion was significantly more abundant than the m/z 1154^{1-} ion (Figure 5.3) while the m/z 1154^{1-} ion became more abundant in the spectra of the fucosylated glycan (Figure 5.4). The Y_1 ion corresponding to the loss of a single GlcNAc was seen in high abundance at

m/z 658^{2-} (Figure 5.3). While the loss of a GlcNAc residue was shown occurring on the 3-arm, cleavage of the 6-arm GlcNAc or the bisecting GlcNAc was equally possible.

The m/z 893.3^{2-}, m/z 966.4^{2-}, and m/z 1039.3^{2-} glycans were identified as bi-antennary N-linked glycans containing varying degrees of fucosylation (Figures 5.5, 5.6, 5.7, 5.8). Since two isomers of the m/z 966.4^{2-} glycan were isolated, the early eluting isomer was labelled isomer 1 and the later eluting isomer labelled isomer 2. Core fucosylation was found to be present in the m/z 966.4^{2-} isomer 1 and m/z 1039.3^{2-} glycans, and absent in the m/z 893.3^{2-} and m/z 966.4^{2-} isomer 2 glycans based on the core related diagnostic ions (Table 5.1, Core). All four glycans exhibited prominent B m/z 510^{1-} and F m/z 570^{1-} ions diagnostic of terminal fucosylation. Fucose substitution on the GlcNAc residue was confirmed by the production of the C$_2$ m/z 528^{1-} ion (Harvey et al., 2008). While fucose substitution on the Gal would still produce the C$_2$ m/z 528^{1-} ion, a C$_1$ m/z 325^{1-} ion would be produced but was absent from all spectra.

Fucose substitution on the 3-arm was confirmed in the m/z 893.3^{2-} and m/z 966.4^{2-} isomer 2 glycans indirectly by the production of D and D-18 ions at m/z 688^{2-} and m/z 670^{2-} respectively corresponding to the presence of Gal and GlcNAc attached to the 6-arm. As the m/z 893.3^{2-} glycan was not core fucosylated, the 3-arm remained the only position available for fucose substitution. Since the m/z 966.4^{2-} isomer 2 was core fucosylated, the 3-arm was similarly the only position available for fucose substitution since the composition for m/z 966.4^{2-} calculated to contain 2 fucose residues. Fucose substitution on the 6-arm was confirmed in the m/z 966.4^{2-} isomer 1 and m/z 1039.3^{2-} glycans by the production of D and D-18 ions at m/z 834^{2-} and m/z 816^{2-} respectively corresponding to the presence of Gal, GlcNAc and the presence of an additional fucose attached to the 6-arm. The key diagnostic ions that differentiate the two isomers of m/z 966.4^{2-} are boxed in Figures 5.6 and 5.7.
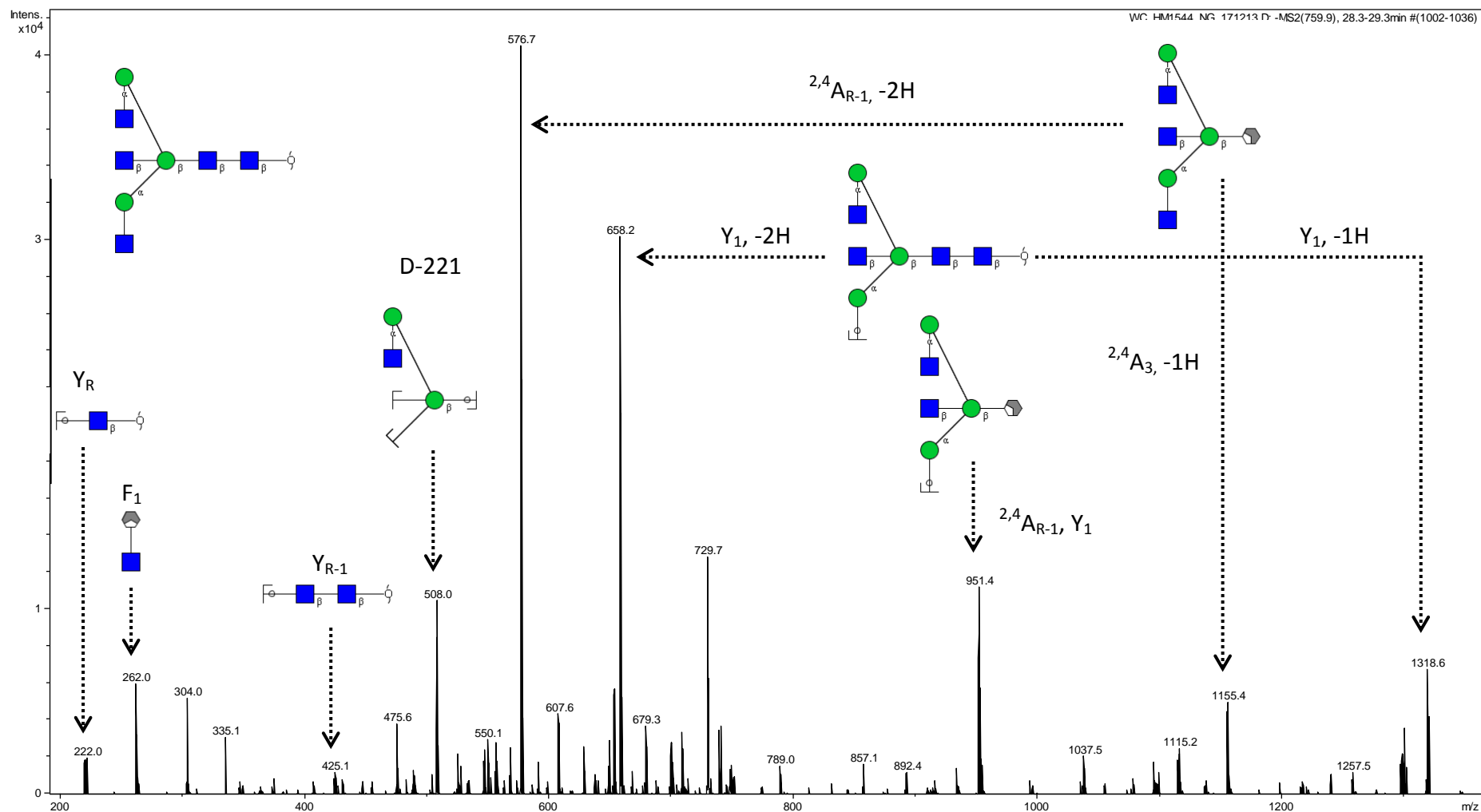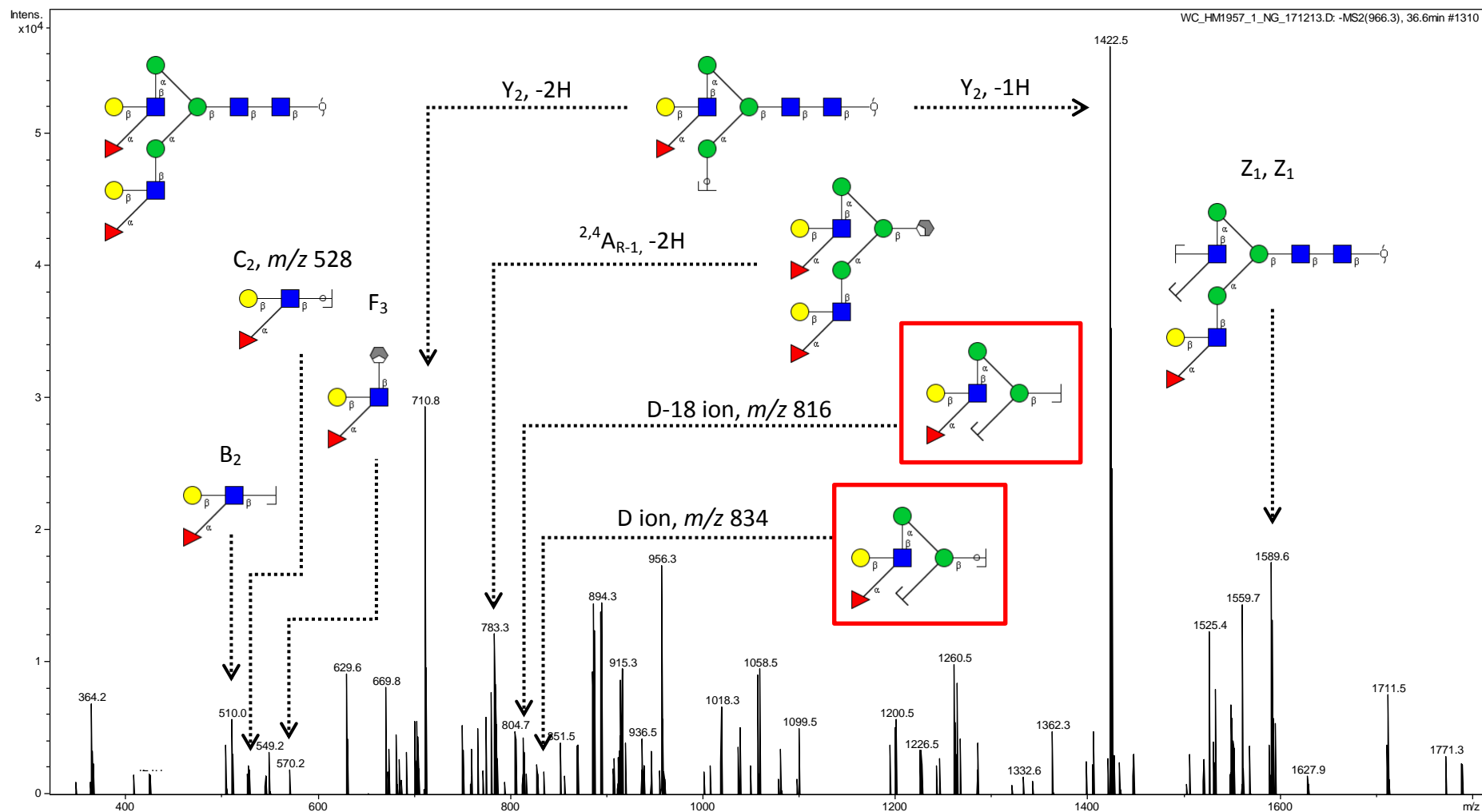
Figure 5.3 – ESI-MS/MS fragmentation spectra of *m/z* 759.9^{2-}. The abundant *m/z* 576^{1-} and *m/z* 1154^{2-} ion indicate $^{2,4}A_3$ cleavage of the core GlcNac attached to the branching mannose. Z cleavage produced ions *m/z* 658^{2-} *m/z* 1318^{1-} corresponding to the loss of a terminal GlcNAc. $^{2,4}A_{R-1}$ cleavage combined with Z cleavage also produced an ion seen as *m/z* 951^{1-} corresponding to the loss of the core GlcNAc and a terminal GlcNAc.

97
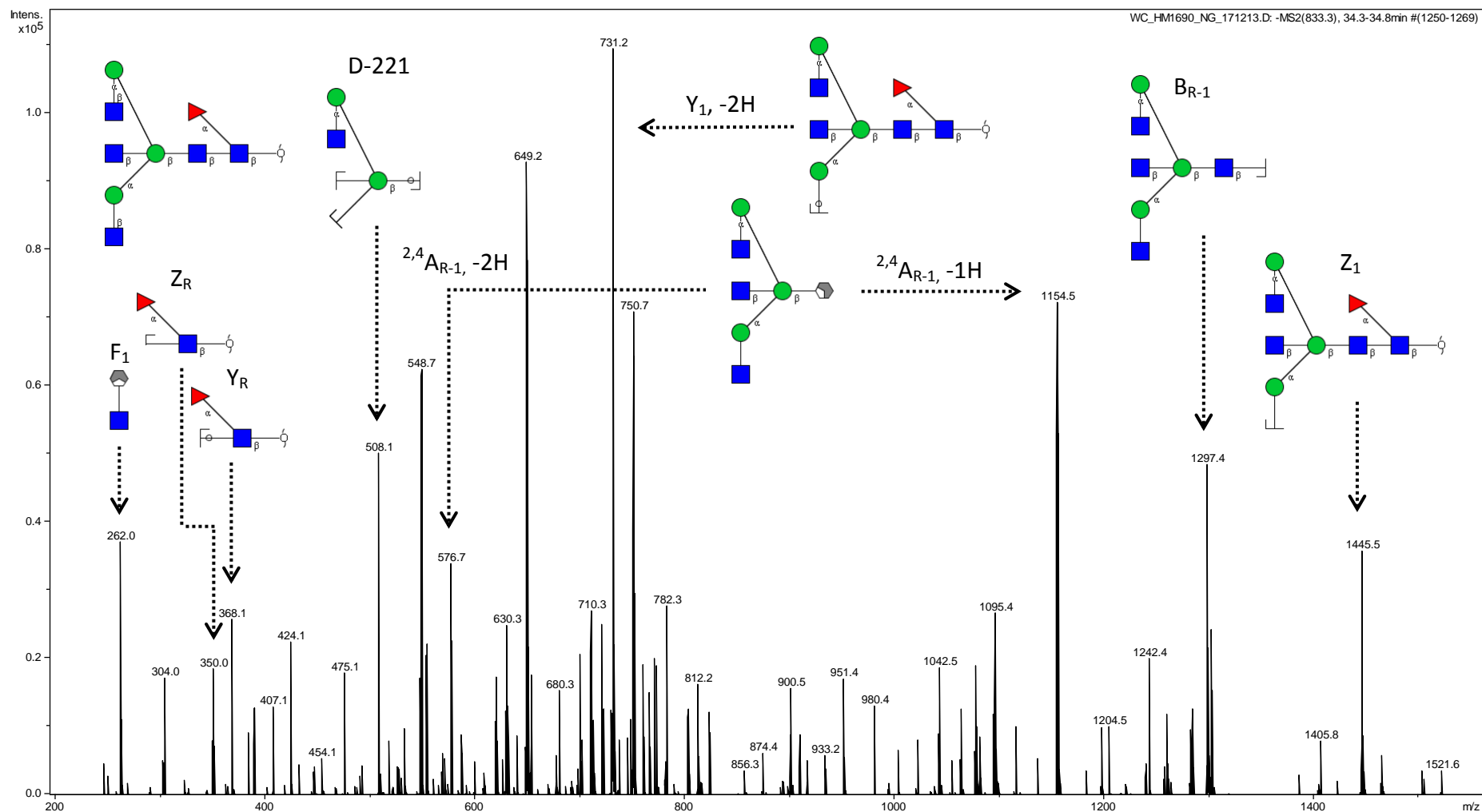
Figure 5.4 – ESI-MS/MS fragmentation spectra of *m/z* 832.8^{2-}. A similar fragmentation spectra as Figure 5.3 was observed except the abundance of the *m/z* 1154^{1-} ion dramatically increased. Core fucosylation was confirmed by the *m/z* 350 and *m/z* 368 diagnostic ions and the resulting B$_{R-1}$ ion can be seen in high abundance at *m/z* 1297. Glycosidic cleavage of a terminal GlcNAc produced an abundant Y ion (*m/z* 731^{2-}) and Z ion (*m/z* 1445^{1-}).
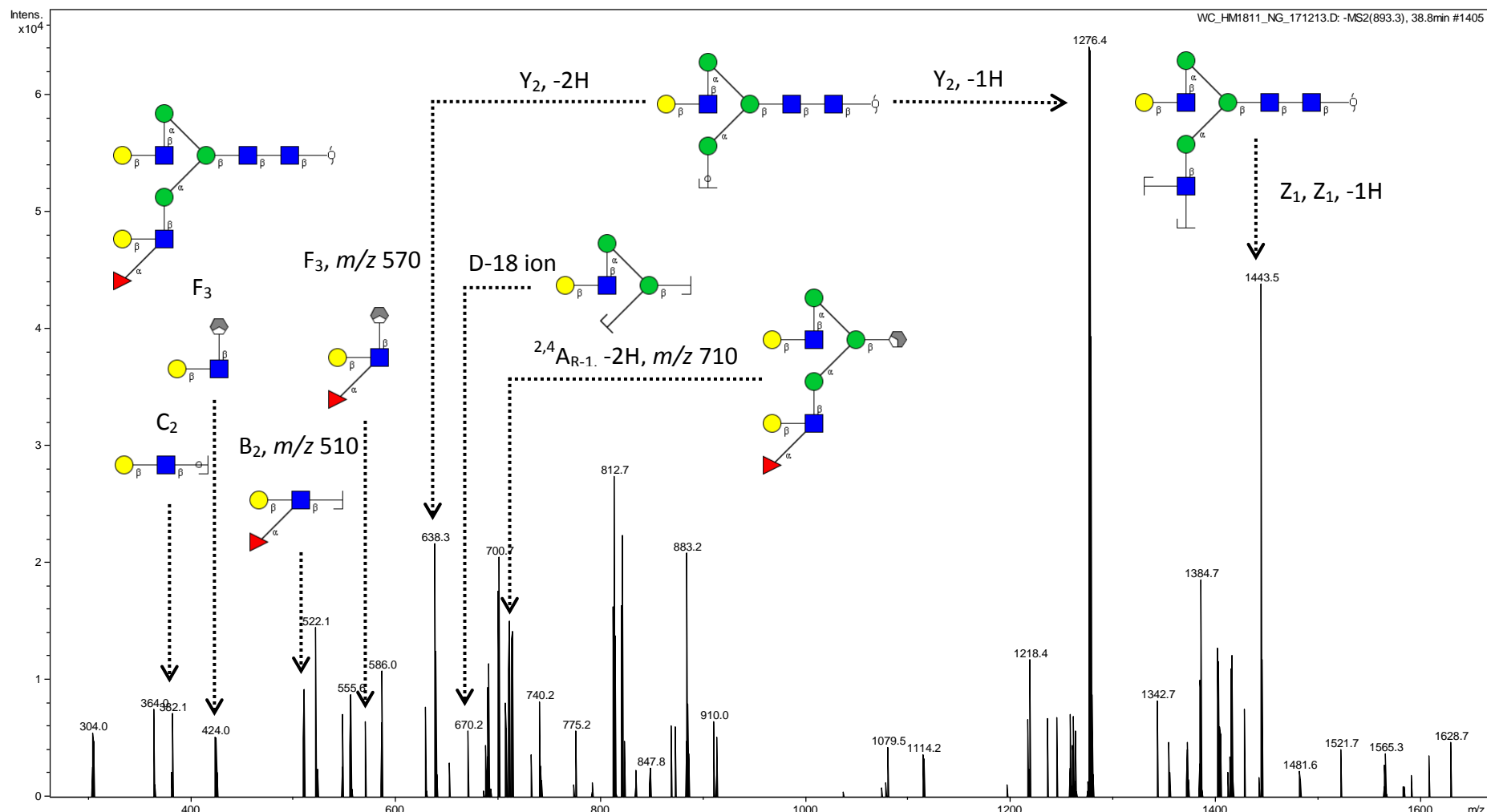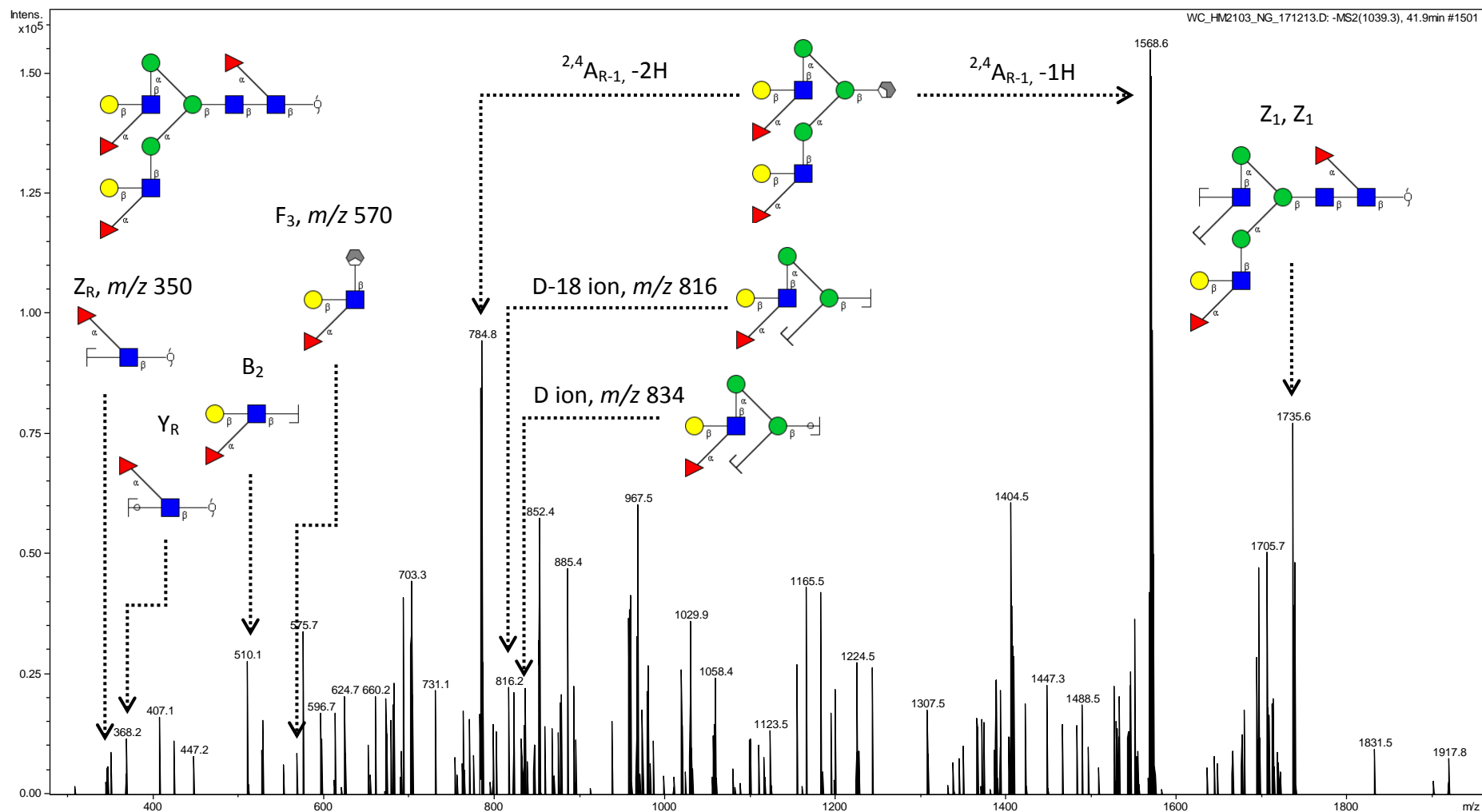
Figure 5.5 – ESI-MS/MS fragmentation spectra of *m/z* 893.3²⁻. Terminal fucosylation was identified by the B and F diagnostic ions *m/z* 510 and *m/z* 570 respectively. Fucosylation of the 3-arm was indicated by the D (*m/z* 688, present but not shown) and D-18 ions (*m/z* 670) which corresponded to a Gal-GlcNAc attached to the 6-arm, leaving the 3-arm as the only remaining point of substitution since there was no sign of core fucosylation. The Y₂ ions corresponding to the loss of the diagnostic B ion (*m/z* 510) can be seen at *m/z* 638²⁻ and *m/z* 1276¹⁻.

Figure 5.6 – ESI-MS/MS fragmentation spectra of *m/z* 966.4^{2-} isomer 1. Terminal fucosylation was identified by the B and F diagnostic ions *m/z* 510 and *m/z* 570 respectively. Key ions that differentiated the two isomers of *m/z* 966.4^{2-} are boxed. The composition of the 6-arm was identified as Gal-Fuc-GlcNAc by the D and D-18 ions (*m/z* 834 and *m/z* 816) with the remaining fucose substituted on the 3-arm.
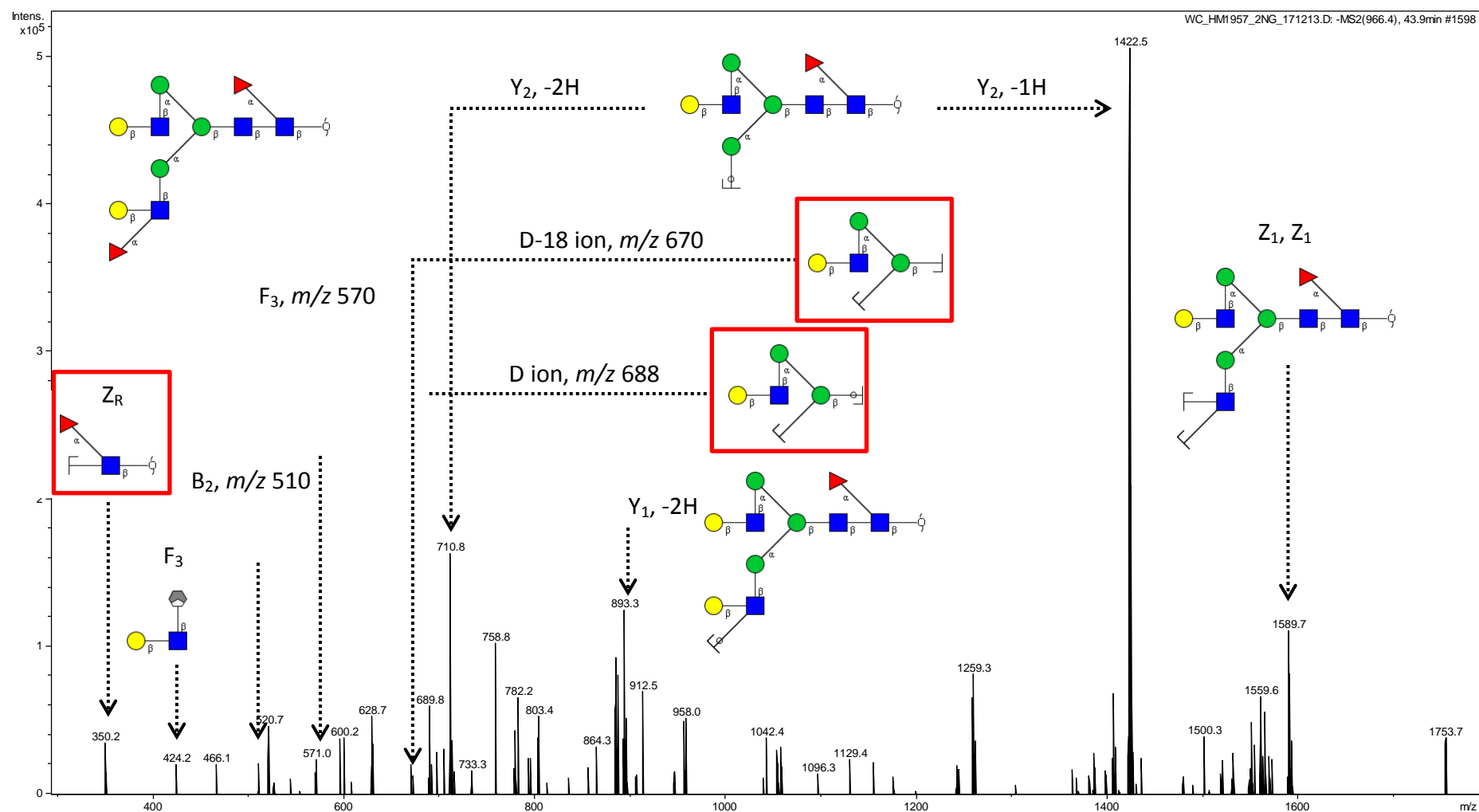
Figure 5.7 – ESI-MS/MS fragmentation spectra of *m/z* 966.4^{2-} isomer 2. Core fucosylation was confirmed by the *m/z* 350 diagnostic ion. Key ions that differentiated the two isomers of *m/z* 966.4^{2-} are boxed. Terminal fucosylation was confirmed by the B and F diagnostic ions *m/z* 510 and *m/z* 570 respectively. Fucosylation of the 3-arm was indirectly confirmed by the D and D-18 ions (*m/z* 688 and *m/z* 670)

Figure 5.8 – ESI-MS/MS fragmentation spectra of *m/z* 1039.3²⁻. Core fucosylation was confirmed by the presence of both the *m/z* 350 and *m/z* 368 diagnostic ions. Terminal fucosylation was confirmed by the B and F diagnostic ions *m/z* 510 and *m/z* 570 respectively with fucosylation present on the 6-arm confirmed by the D and D-18 ions (*m/z* 834 and *m/z* 816)

Fucose substitution via α1–3 linkage was confirmed by exoglycosidase digestion with β(1-4)-galactosidase, which specifically cleaves β1 − 4 linked Gal, and α(1-3,4)-fucosidase, which cleaves both α1 − 3 and α1 − 4 linked Fuc (Figure 5.9).
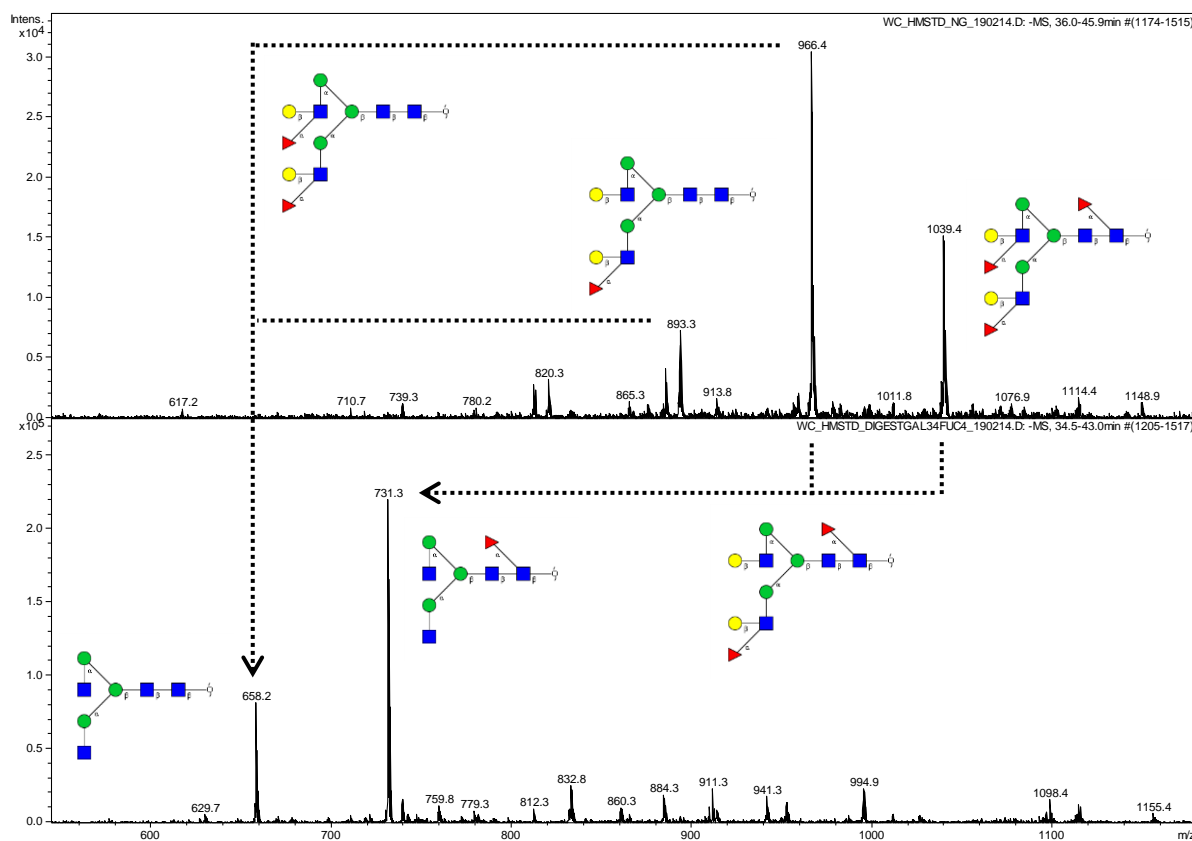
Figure 5.9 – Comparison of the average mass spectrum of the mixture of the six *N*-linked glycans before (top) and after (bottom) exoglycosidase digestion to confirm the linkage of terminal Gal and Fuc residues. Exoglycosidase digestion resulted in the formation of *m/z* 731^{2-} or *m/z* 658^{2-} corresponding to the core fucosylated or non-core fucosylated *N*-linked glycans following loss of β1 − 4 linked Gal and α1 − 3 linked Fuc.

As a prerequisite for fucosidase digestion the terminal galactose must first be removed (Royle et al., 2007). Since the digestion of Gal and Fuc residues was successful, the linkage of Gal was confirmed as β1 − 4 (Figure 5.9). Because the 4 position of GlcNAc was substituted by Gal, the linkage of Fuc was confirmed as α1 − 3 considering the specificity of the fucosidase (both α1 − 3, 4) and assuming the biosynthetic rules for fucosylation (Section 1.1.4).

5.3.2 Establishment of a Retention Time Library

In the work presented in Chapter 4, *N*-linked glycans released from human skim milk proteins were separated on a preparative scale by PGC-HPLC and glycans were detected by offline

MALDI-TOF MS analysis of collected fractions (Section 4.2.3). Based on the masses obtained from the MALDI-TOF spectra of each fraction, low molecular weight glycans generally eluted early followed by glycans of increasing molecular weight. However, there were exceptions to this behavior with isomeric structures of the same molecular weight eluting at considerably different times, for example, two isomers of m/z 1957^{1+} were separated and collected at different retention times with the first isomer eluting at 56 − 57 minutes and the second isomer eluting at 62 − 63 minutes (Section 4.2.3). Additionally, some cases of higher molecular weight glycans eluting before glycans of lower molecular weight, for example m/z 2103^{+} eluted earlier than one of m/z 1957^{+} isomers (Figure 4.5). To further examine the elution pattern of the glycans, the six human milk N-linked glycans purified by PGC-HPLC (Chapter 4) were recombined and analysed as a mixture by PGC-LC-ESI-MS (Figure 5.10). The extracted ion chromatogram (EIC) allows m/z values to be selected from a mass chromatogram and the EICs were used to compare specific structures or isomers. By relating structural epitopes to the elution pattern of the purified N-linked glycans on PGC-LC-ESI-MS, a retention time library was established.

It was clearly evident that certain glycans were more abundant or had different ionisation efficiencies than others in the mixture. The N-linked glycans consistently maintained the same retention time when analysed individually or as a mixture, which was suggestive of their suitability as internal standards in the analysis of unknown samples.
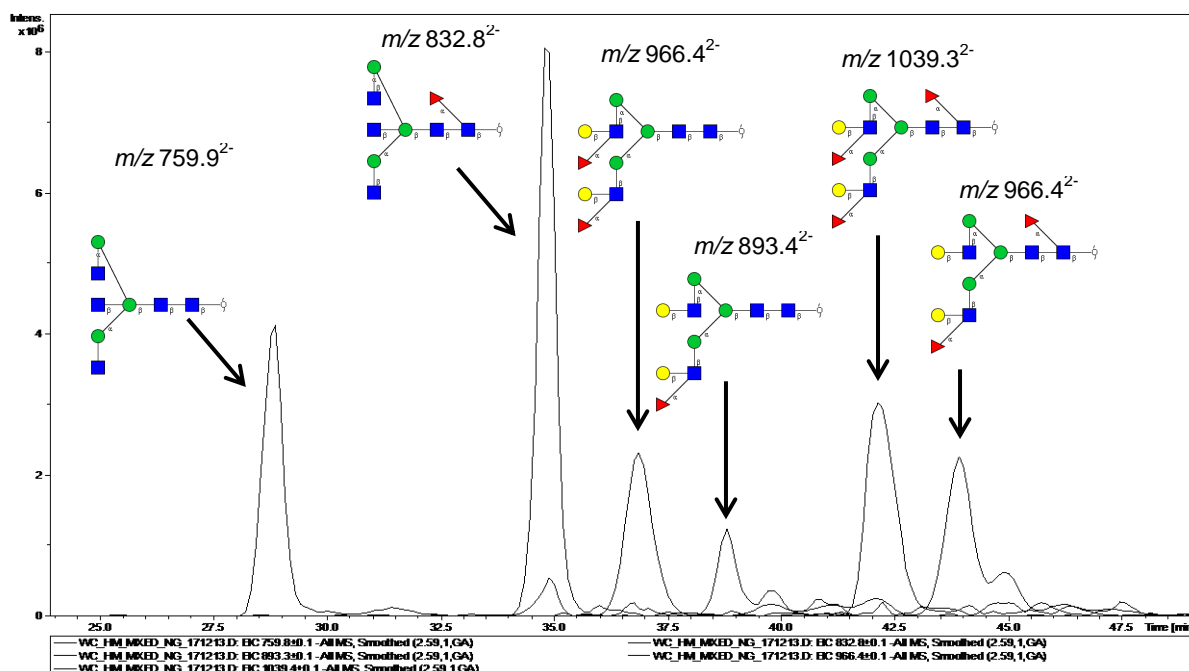
Figure 5.10 – The average EIC of six *N*-linked glycans isolated from human milk and purified by PGC-HPLC (Chapter 4).

The structural characterisation of the purified *N*-linked glycans by ESI-MS/MS (Section 5.3.3) provided an explanation of the elution pattern observed by MALDI-TOF MS analysis of the PGC-HPLC fractions in Chapter 4 and the EIC from the PGC-LC-ESI-MS in the work described in this chapter. For instance, the presence of an $\alpha 1 - 6$ linked core fucose appeared to increase retention on PGC. This was most evident when comparing *m/z* 759.9^{2-} (retention time 28.9 mins) and *m/z* 832.8^{2-} (retention time 34.8 mins), which differed only in the core linked fucose (Figure 5.11). The presence of a core fucose in the case of *m/z* 832.8^{2-} contributed to an increase in retention time of ~6 mins.
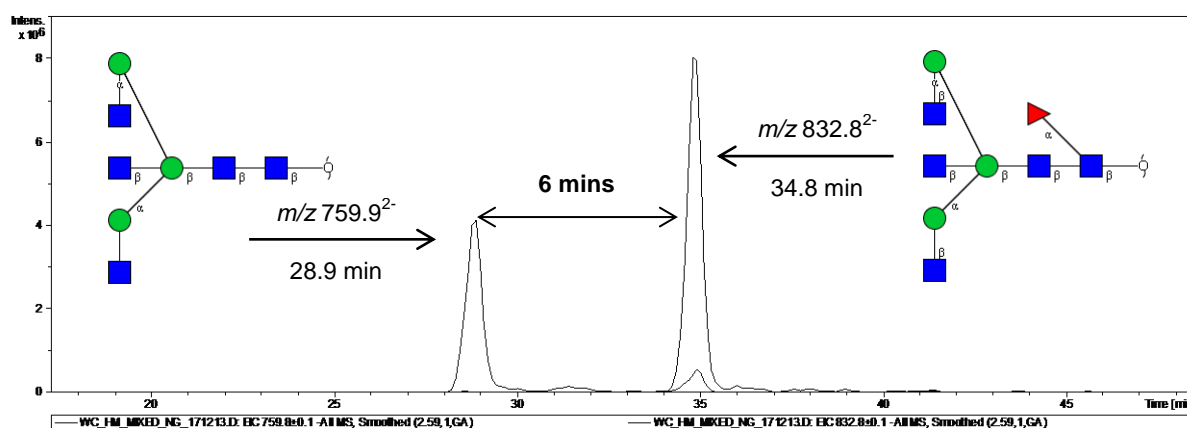
Figure 5.11 – Comparison of the EIC of *m/z* 759.9²⁻ and *m/z* 832.8²⁻ *N*-linked glycans containing a bisecting GlcNAc. The presence of a core linked fucose resulted in a 6 minute increase in retention time.

A similar retention shift was seen when comparing the two isomers of *m/z* 966.4²⁻ (Figure 5.10). Structural analysis of *m/z* 966.4²⁻ by PGC-LC-ESI-MS/MS (Section 5.3.2) identified two isomers of a bi-antennary glycan containing two fucose residues with different positions of substitution. The glycan of *m/z* 966.4²⁻ isomer 1 was found to contain two α1 – 3 linked terminal fucose residues, also known as the Lewis x (Le^x) antigen (Figure 5.10, retention time 36.9 min), while the *m/z* 966.4²⁻ isomer 2 was found to contain a single Le^x antigen on the 3-arm and an α1 – 6 linked core fucose (Figure 5.10, ret time 43.9 min).
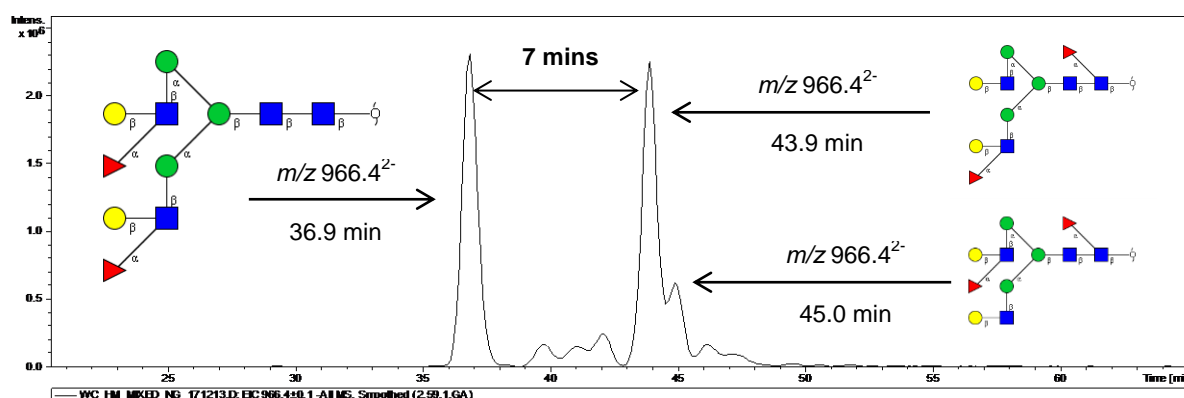
Figure 5.12 – Comparison of the EIC of *m/z* 966.4²⁻ structural isomers, the early eluting isomer contains two terminal fucose residues, while the later eluting isomer contains a core fucose and one terminal fucose.

When the retention times of the two isomers were compared, the core fucosylated *m/z* 966.4²⁻ isomer 2 had an even greater shift in retention time eluting ~7 mins later. Additionally, it was also possible that the presence and position of the Le^X antigens may have influenced retention

106

time. Upon further investigation, the tailing peak (Figure 5.12, retention time 45.0 min) of the later eluting isomer of m/z 966.4^{2-} was identified to contain the Le^{x} antigen substituted on the 6-arm evidenced by a prominent m/z 834^{2-} (Table 5.1, D-ion) and m/z 816^{2-} (Table 5.1, D-18) ions in the fragmentation spectra. This finding indicated that the m/z 966.4^{2-} isomer 2 fraction was in fact a mixture of two structures; the Le^{x} antigen on the 3-arm as the major component (~75%) and the Le^{x} antigen on the 6-arm as the minor component (~25%). The elution pattern observed suggests that substitution of the Le^{x} antigen on the 3- or 6-arm contributes to shifts in retention time; substitution on the 6-arm results in a ~1 min increase in retention time on PGC. The influence of the Le^{x} antigen was further investigated by comparing the EIC of m/z 966.4^{2-} isomer 1 (retention time 36.9 mins) and m/z 893.4^{2-} (retention time 38.4 mins, Figure 5.12).
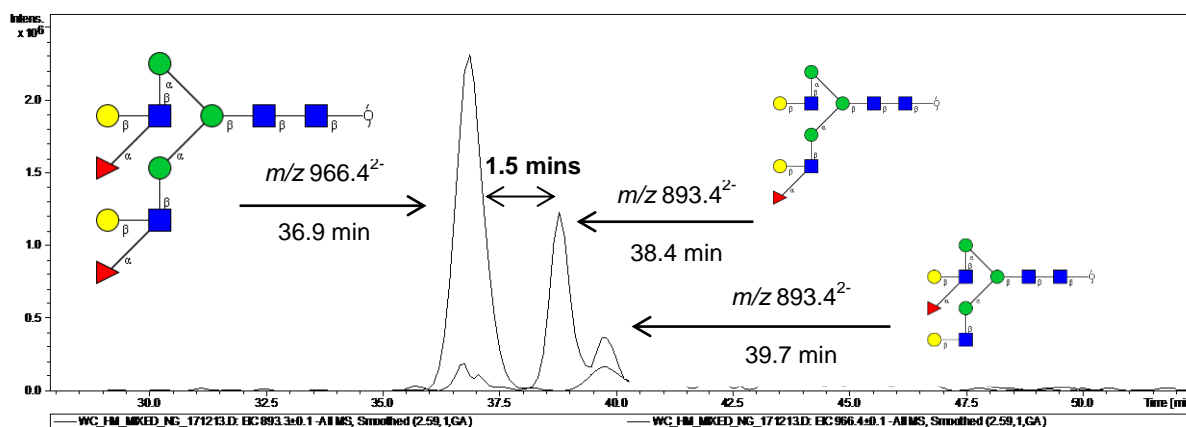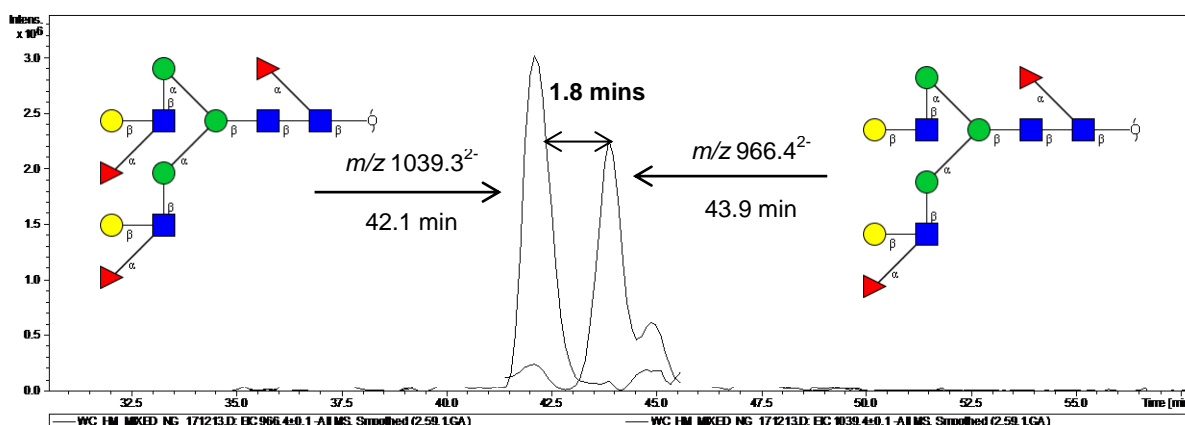

Figure 5.13 – Comparison of the EIC of the early eluting isomer of m/z 966.4^{2-} (left) and m/z 893.4^{2-} (right) N-linked glycans which differ by one terminally substituted fucose.

Based on size alone, the larger m/z 966.4^{2-} isomer 1 was expected to elute after the smaller m/z 893.4^{2-}, however, when actual retention times were compared, the larger glycan of m/z 966.4^{2-} isomer 1 eluted slightly earlier than the smaller m/z 893.4^{2-}. Structural analysis of m/z 893.4^{2-} (Section 5.3.2) identified a bi-antennary glycan containing Le^{x} antigen substituted on the 3-arm. When comparing m/z 966.4^{2-} isomer 1 and m/z 893.4^{2-}, the presence of an extra Le^{x} antigen in the case of m/z 966.4^{2-} contributed to a decrease in retention time by ~1.5 mins. A tailing peak was also observed in the EIC of m/z 893.4^{2-}, which was identified to contain the

Le^x antigen on the 6-arm by a prominent m/z 834^{2-} ion (D-ion) in the fragmentation spectra, indicating that the m/z 893.4^{2-} fraction was also a mixture of two structures; the Le^x antigen on the 3-arm as the major component (~70%) and the Le^x antigen on the 6-arm as the minor component (~30%). The slight (~1 min) increase in retention of the tailing peak was consistent with the behavior observed in the comparison of the two m/z 966.4^{2-} isomers, which suggests that substitution of the Le^x antigen on the 6-arm of a bi-antennary glycan results in a ~1 min increase in retention on PGC.

The observed contribution of Le^x antigen to decreased retention time on PGC was also seen when comparing the later eluting isomer m/z 966.4^{2-} (retention time 43.9 mins) with m/z 1039.3^{2-} (retention time 42.1 mins, Figure 5.13). Structural analysis of m/z 1039.3^{2-} (Section 5.3.2) identified a bi-antennary glycan containing a single Le^x antigen on each arm and an α1 – 6 linked core fucose (Figure 5.13). Based on size, the smaller m/z 966.4^{2-} was expected to elute before the larger m/z 1039.3^{2-}, however, when actual retention times were compared, the larger m/z 1039.3^{2-} eluted earlier than the smaller m/z 966.4^{2-}.

Figure 5.14 – Comparison of the EIC of the later eluting isomer of *m/z* 966.4^{2-} and *m/z* 1039.3^{2-} *N*-linked glycans containing a core substituted fucose that differ by one terminally substituted fucose.

When comparing m/z 1039.3^{2-} with the later eluting isomer of m/z 966.4^{2-} the presence of an extra Le^x antigen on the 6-arm resulted in a decrease in retention time of ~ 2 mins. The work described in this section supports the use of PGC in the separation of closely related *N*-linked

glycans. PGC was capable of separating isomeric structures and the substitution of α1 − 6 linked core fucose and the Le^x antigen contributed to increased and decreased retention on respectively. Furthermore, substitution of the Le^x antigen on the 6-arm was found to cause a slight (~1 min) increase in retention compared to substitution on the 3-arm. The availability of the purified glycans allowed these retention rules to be defined.

5.3.3 Characterisation of the *N*-linked Glycoprofile of Human Milk Whey Proteins

In order to establish the relative abundance of the six *N*-linked glycans isolated as standards in this work, within the entire *N*-linked glycoprofile of human milk proteins, the PGC-LC-ESI-MS data from the work described in Section 3.2.7 was re-examined (Figure 5.15). Only glycan masses with MS/MS data were used and glycan compositions were calculated using GlycoMod. Relative abundance was obtained by integration of peaks from the EIC of glycan masses and represented as a percentage of the summed total abundance of all glycans detected. The six neutral *N*-linked glycans purified by preparative scale PGC-HPLC were found to be among the most abundant species within the entire *N*-linked glycan profile of the milk proteins with the exception of the truncated bi-antennary glycan with a bisecting GlcNAc (HexNAc)3+(Man)3(GlcNAc)2 (Figure 5.15). The majority of *N*-linked milk glycans attached to proteins were neutral glycans based on composition with ~78% considered neutral while the remaining ~22% were acidic glycans containing at least one NeuAc residue. Of the neutral glycans, ~31% were high mannose type and ~69% were complex type. Fucosylation was a prominent feature of the human milk protein *N*-linked glycans with ~53% containing fucosylation, ~19% containing both fucosylation and sialylation (NeuAc), and only ~2% found to contain only sialylation.
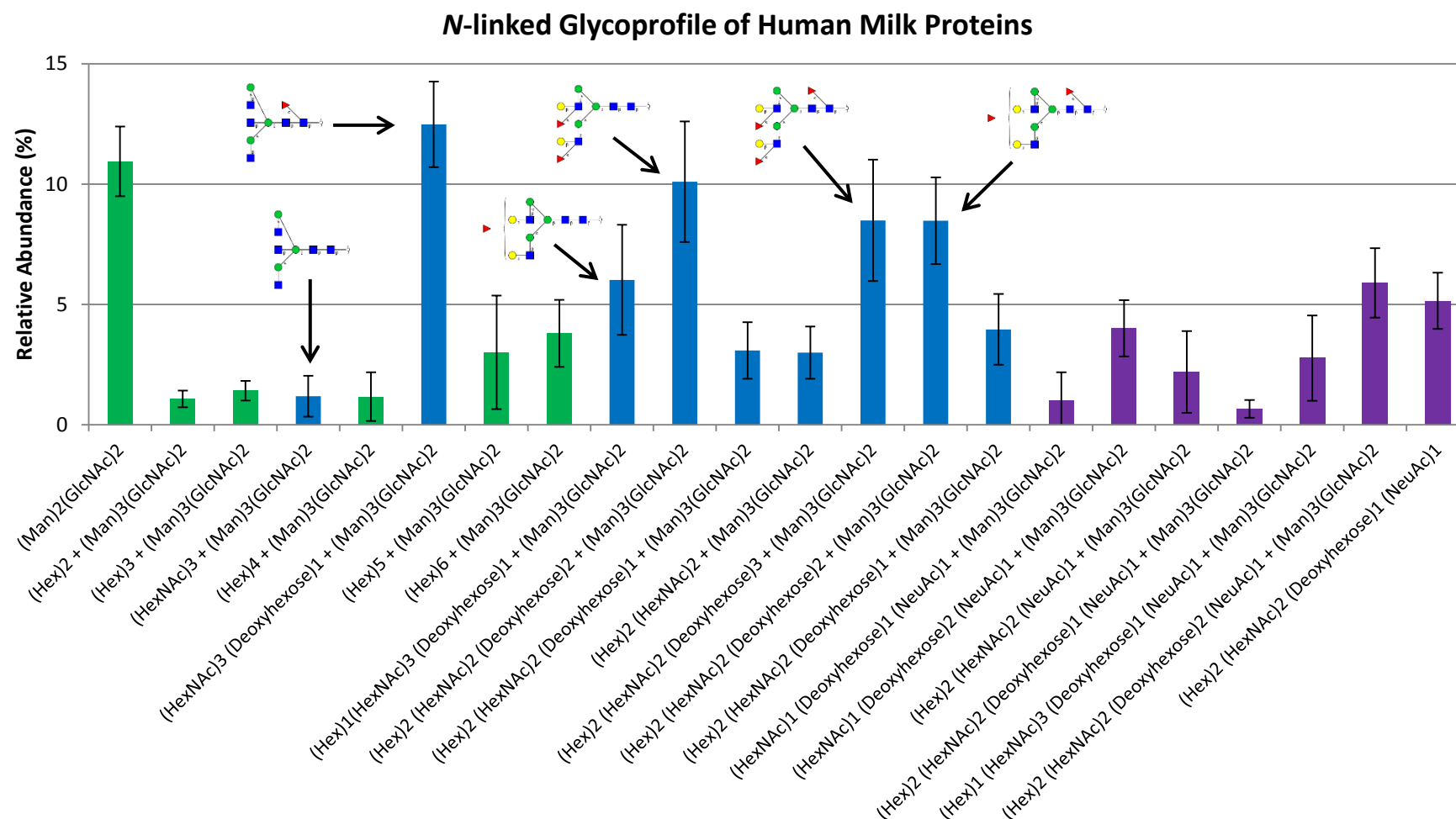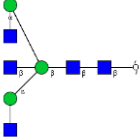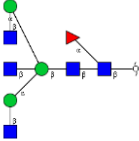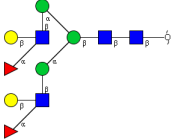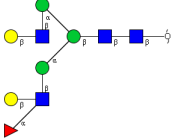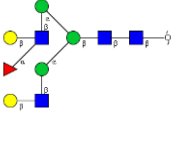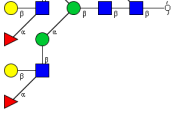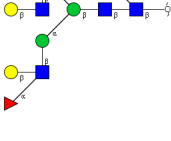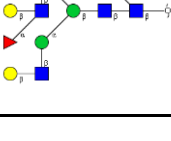
Figure 5.15 – Relative abundance of the human milk protein *N*-linked glycans as determined by PGC-ESI-LC-MS data. Glycan identities are given as compositions on the horizontal axis, relative abundance is given on the vertical axis calculated as a percentage of the summed total of all *N*-linked glycans. Three types of *N*-linked glycan are represented by colour; Green: high mannose; Blue: neutral complex; Purple; acidic (containing NeuAc) complex. The structures of the six purified neutral *N*-linked glycans are shown with arrows indicating the composition each represents within the glycoprofile.

5.4 Chapter Summary

Having separated and collected purified *N*-linked glycans released from human milk proteins on a preparative scale, complete structural characterisation was achieved using the established PGC-LC-ESI-MS method in order to validate their use as glycan standards (Jensen et al., 2012). Single stage MS data was analysed using GlycoMod to calculate potential glycan compositions and the MALDI-TOF data validated by ESI-MS. Tandem MS fragmentation data was processed manually aided by the use of GlycoWorkbench and key diagnostic ions. Major structural features such as the sequence of the 3- and 6-arms were identified. Linkage positions predicted as terminal $\beta 1 - 4$ Gal and $\alpha 1 - 3$ Fuc residues of the Lewis x epitopes were confirmed by exoglycosidase digestion with β1-4-galactosidase and α1-3,4-fucosidase. Following MS/MS structural characterisation, a retention time library was established using the six purified *N*-linked glycans released from human milk proteins, which revealed two additional structures with unique retention times as shown below (Table 5.3). LC analysis of the eight structures illustrated that the $\alpha 1 - 6$ linked core fucose and both the presence and position of the Lewis x epitope had an influence on retention time on PGC. Characterisation of the compositions, structures and relative abundance of all the *N*-linked glycans released from human milk proteins revealed that the *N*-linked glycans purified for use as standards were highly representative of the most abundant structures.

Table 5.3 – The *N*-linked glycans purified by PGC-HPLC and structurally characterised by a combination of MALD-TOF MS and PGC-LC-ESI-MS in order of elution.

MALDI-TOF [M + Na]$^{1+}$ *m/z*	ESI-MS [M – 2H]$^{2-}$ *m/z*	PGC-LC-MS Retention time (mins)	Structure
1544	759.9	28.9	
1690	832.8	34.8	
1957	966.4	36.9	
1811	893.3	38.4	
1811	893.3	39.7	
2103	1039.3	42.1	
1957	966.4	43.9	
1957	966.4	45.0	

Chapter 6 – Application of Human Milk *N*-linked Glycans to an *in vitro* Bacterial Binding Assay

6.1 Introduction

The availability of preparative scale quantities of *N*-linked glycans potentially allows for other applications in addition to use as chromatography and MS standards. The involvement of glycoproteins in host-pathogen interactions has been of increasing interest and a number of *in vitro* techniques have been developed to investigate these interactions; however, very few studies have focused specifically on conjugated glycans as opposed to free oligosaccharides or the entire glycoprotein in host-pathogen interactions. Recent interest in the bioactivity of human milk has implicated a number of free oligosaccharides and human milk glycoproteins in the protection of breastfed neonates against pathogenic infection by virtue of their ability to inhibit the adherence of pathogens to the gut (as reviewed in Peterson et al., 2013). Therefore, the work described in this chapter involved the application of the *N*-linked glycans released from human milk glycoproteins (Chapter 3) in an established in-house assay designed to measure bacterial adhesion to human intestinal cells, with the aim of investigating the involvement of the human milk *N*-linked glycans in the binding interactions.

6.1.1 The Anti-adhesive Activity of Human Milk

Human infants are born with an undeveloped immune system rendering them highly susceptible to a vast range of pathogenic infections immediately after birth, as they are simply unable to mount an immune response (Newburg et al., 2005). It has been well established that breastfed infants have significantly lower incidence of acute gastrointestinal disease when compared to non-breast fed infants (Grulee et al., 1935; Morrow and Rangel, 2004; Newburg, 2009). Protection against gut pathogens has been partly attributed to the free oligosaccharides, which competitively bind to bacteria, thus reducing their adhesion to the glycan receptors in the gastrointestinal tract (Liu and Newburg, 2013; Newburg, 2009; Peterson et al., 2013). The

theory advocating the anti-adhesive potential of milk glycans was first proposed for the human milk oligosaccharides (HMOs) (Bode, 2012; Newburg et al., 1990), but has since been extended to the glycans of human milk glycoconjugates. Since the biosynthesis of conjugated glycans on membrane and secreted proteins occurs via the same pathway (Section 1.1.4), human milk glycoproteins have a high potential to possess glycans that resemble those found on the surface of human epithelial cells.

The onset of antibiotic resistant strains of bacteria has seen growing interest in the bioactivity of human milk proteins, which could potentially lead to glycan based anti-adhesive therapeutics (Sharon, 2006). Anti-adhesive therapies provide a number of advantages over typical antibiotic treatments. Oligosaccharides of natural origin may have a broad spectrum of activity and are less likely to be toxic or immunogenic (Sharon, 2006). Since anti-adhesives are neither bacteriocidal or bacteriostatic, the lack of sustained selective pressures reduces the chance of resistant strains emerging (Kelly and Younson, 2000).

Pathogen-glycan interactions are mediated by glycan binding proteins (GBPs) such as the bacterial lectins or the influenza virus haemagglutinin (Karlsson, 1995). Research on the anti-adhesive activity of HMOs has identified terminal fucose and NeuAc as prominent binding epitopes for pathogen GBPs (reviewed by Bode, 2012). A number of bacterial lectins have been shown to recognise epitopes such as the Lewis b (Fucα1 – 2Galβ1 – 3[Fuc α1 – 4]GlcNAc), and 3-sialyllactose (NeuAcα2 – 3Galβ1 – 4Glc), which are common structures found in HMOs (Sharon, 2006). Highly abundant human milk glycoproteins such as lactoferrin, secretory immunoglobulin A (sIgA), and mucins, have also been shown to inhibit the adhesion of pathogens to human epithelial cells. However, the majority of studies on milk glycoproteins have simply identified the potential anti-adhesive activity of the glycoprotein as a whole (reviewed by Peterson et al., 2013) and very few studies have specifically investigated the anti-adhesive activity of the glycan component (Table 6.1).

Lactoferrin and secretory immunoglobulin A (sIgA) are the main glycoproteins in human milk that have been shown to possess anti-adhesive activity against pathogens that is influenced by *N*-linked glycosylation (Table 6.1). The *N*-linked glycosylation of human lactoferrin has been extensively studied and shown to contain mainly complex type glycans with bi-, tri- and tetra-antennary *N*-acetyllactosamine (Galβ1 – 4GlcNAc) branches substituted with terminal fucose (Le^x epitope) and NeuAc residues (Barboza et al., 2012; Samyn-Petit et al., 2003). Comparatively, sIgA contains predominantly truncated (lacking Gal) bi-antennary structures substituted with a core fucose and bisecting GlcNAc. In addition, ~12% of sIgA *N*-linked glycans are high mannose type containing between five and nine mannose residues (Man$_{5-9}$GlcNAc$_2$) (Arnold et al., 2007). The *in vitro* studies performed on lactoferrin and sIgA (Table 6.1) provide substantial evidence for the involvement of *N*-linked glycosylation in the anti-adhesive protection against pathogens.

Table 6.1 – List of human milk *N*-linked glycoproteins and associated glycan epitopes that have been shown to inhibit the adhesion of pathogens to human cells, or have high affinity for pathogen GBPs, adapted from (Peterson et al., 2013).

Human Glycoproteins and Glycan Epitopes Involved in Adhesion	Pathogen/ Toxin	Binding Target	Reference
Bile salt stimulated lipase			
Lewis x epitope	Human immunodeficiency virus (HIV-1)	Human CD4 + T cells	(Naarding et al., 2006)
Lactoferrin			
N-linked glycans	*Listeria monocytogenes*	Human Caco-2 cells (epithelial colorectal adenocarcinoma)	(Barboza et al., 2012)
Terminal galactose	*Salmonella enterica* Heidelberg	Human Caco-2 cells (epithelial colorectal adenocarcinoma)	(Barboza et al., 2012)
N-linked glycans	*Salmonella enterica* enteridis	Human Caco-2 cells (epithelial colorectal adenocarcinoma)	(Barboza et al., 2012)
Terminal fucose	*Salmonella typhimurium*	Human Caco-2 cells (epithelial colorectal adenocarcinoma)	(Barboza et al., 2012)
Terminal fucose, Lewis x epitope	*Pseudomonas aeruginosa*	Lactoferrin showed high affinity for *P. aeruginosa* lectin PA-IIL	(Lesman-Movshovich et al., 2003)
Secretory Immunoglobulin A (sIgA)			
N- and *O*-linked Glycans	*Clostridium difficule* Toxin A	Hamster intestinal brush border membranes	(Rolfe and Song, 1995)
Terminal fucose	Enteropathogenic *E. coli*	Human HEp-2 cells (epithelial laryngeal carcinoma)	(Cravioto et al., 1991)
Terminal NeuAc	S-fimbriated *E. coli*	Human buccal epithelial cells	(Schroten et al., 1993)
High mannose	Type 1 fimbriated *E. coli*	Human HT-29 cells (epithelial colorectal adenocarcinoma)	(Wold et al., 1990)
Terminal fucose	*Helicobacter pylori*	Human epithelial gastric tissue mucosa cells	(Falk et al., 1993)

6.1.2 Measurement of the Anti-adhesive Activity of Human Milk Oligosaccharides and Glycoproteins

A number of *in vitro* techniques have been developed to study bacterial adherence to human cells. Traditional techniques involved the immobilisation of human epithelial cells (usually from cultured cell lines) onto multi-welled plates followed by the application of bacteria (Nizet et al., 1998). Adherence of the bacteria was determined by plate count techniques in which unbound pathogens were removed, plated onto agar, and colony forming units (CFU) counted. The effect of pre-incubation of the bacteria with potential inhibitors of bacterial binding was determined by subtracting the CFU of the inhibited pathogen from the CFU of an uninhibited control (Nizet et al., 1998).

The development of fluorescent bacterial stains such as SYBR Green has enabled rapid quantitation of bacterial adherence in cell based assays via fluorescent plate reader without the need for subsequent plate colony counting (Grootaert et al., 2011). Enzymatic techniques to investigate the anti-adhesive activity of specific glycan epitopes have also been developed to confirm the involvement of terminal glycan epitopes (Barboza et al., 2012). Enzymatic removal of terminal fucose and NeuAc residues from glycoproteins has been shown to influence adhesion of bacteria to human epithelial cells in cell based assays (Barboza et al., 2012).

In this work, anti-adhesive activity was assessed using an established "PVDF binding assay" developed in-house. The assay was performed on 96 well plates that have a polyvinylidene fluoride (PVDF) membrane at the bottom of the wells. In contrast to whole cell based assays, the PVDF binding assay involves the immobilisation of human intestinal epithelial cell membrane proteins to the PVDF membrane, allowing the measurement of pathogen-membrane protein binding and avoiding the delicate handling procedures required for monolayer live cells assays (Kelly and Younson, 2000). Membrane proteins from LS174T, a

mucin producing human intestinal cell line (Baldwin et al., 1976), were bound to the PVDF plate to model the environment of the human gut. Two bacterial strains were tested for their adherence to the intestinal membrane proteins. The enterohemorrhagic *Escherichia coli* O111:NM (ATCC 43887) is a Gram-negative non-motile opportunistic gut pathogen that produces the Shiga toxin and causes food poisoning and diarrhoea (Xia et al., 2010). The second bacterial strain was *Lactobacillus rhamnosus* (ATCC 53103), a Gram-positive lactic acid producing bacterium that is part of the commensal microbial community of the gut (Morita et al., 2009; Rolfe, 2000).

6.2 Materials and Methods

6.2.1 Cultivation of LS174T human intestinal cells

LS174T intestinal epithelial cells were kindly provided by Professor Hazel Mitchel, School of Biotechnology and Biomolecular Sciences, University of New South Wales. The cells were stored in DMSO at -130 °C. Prior to cultivation, the cells were thawed and re-suspended in culture medium. Cells were cultured in both 200 mL culture flasks (Thermo Scientific) and 100 mL culture plates (Thermo Scientific) in RPMI 1640 medium (Gibco) supplemented with 2 mM Glutamine (Gibco), 10% fetal bovine serum (FBS, Gibco), 100 U/mL penicillin and 100 μg/mL streptomycin (Gibco). Culture media and phosphate buffered saline (PBS) were preheated to 37 °C before use. Cells were grown at 37 °C in a humidified incubator with 5% CO_2. Cells cultured in flasks were passaged and divided into a new culture flask and culture dish every 2-3 days. To passage the cells, the culture medium was removed and the cells were PBS twice. Cells were liberated from flasks by adding trypsin digestion in 0.01 M phosphate buffer, pH 7.2, containing 0.15 M NaCl (PBS) and incubating at 37 °C for 5 minutes. The liberated cells were then divided between a new culture flask and culture plate and supplemented with fresh culture medium. Cells were grown in culture plates until 60% confluency. To harvest cells, the culture medium was removed and the cells washed twice with 4 mL PBS. The cells were then scraped off the dish by spatula and collected in a 2 mL Eppendorf tube, centrifuged gently at 1500 g for 3 minutes and the PBS was removed. The cell pellets were stored at -20 °C until further use.

6.2.2 Extraction of LS174T Membrane Proteins

Total membrane proteins were extracted from LS174T cells using a differential detergent method (Ramsby et al., 1994). LS174T cells were re-suspended by pipetting in 1 mL of ice cold digitonin buffer (0.01% digitonin, 10 mM piperazine-*N*,*N*'-bis(ethanesulfonic acid, PIPES), pH 6.8, 300 mM sucrose, 100 mM NaCl, 3 mM MgCl, 5 mM EDTA, 1.2 mM

phenylmethylsulfonyl fluoride (PMSF) protease inhibitor). The cells were incubated on ice with gentle agitation on a shaker (~450 rpm) for 30 mins then centrifuged at 480 g, 4 °C for 10 mins and the supernatant removed to obtain the cytosolic protein fraction (C1). The cells were re-suspended in 500 μL of ice cold digitonin buffer and incubated on ice with gentle agitation on a shaker (~450 rpm) for 10 mins then centrifuged at 480 g, 4 °C for 10 mins and the supernatant removed to obtain the cytosolic wash fraction (C2). The previous step was repeated to obtain the cytosolic wash fraction (C3).

The insoluble cell pellet was further extracted with 500 μL ice cold Triton X-100 buffer (0.5% Triton X-100, 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 300 mM sucrose, 100 mM NaCL, 3 mM MgCl, 3 mM EDTA, 1.2 mM PMSF protease inhibitor). The cells were incubated on ice with gentle agitation on a shaker (~450 rpm) for 30 mins then centrifuged at 5000 g, 4°C for 10 mins and the supernatant removed to obtain the membrane protein fraction (M1). The cells were re-suspended in 500 μL of ice cold Triton X-100 buffer and incubated on ice with gentle agitation on a shaker (~450rpm) for 10 mins the centrifuged at 5000 g, 4°C for 10 mins and the supernatant removed to obtain the membrane wash fraction (M2). Fractions were separated on an SDS-PAGE gel (Section 2.2.1, data not shown) to confirm similarities and M1 and M2 containing membrane proteins were combined. Protein concentration was quantified by a plate based Bradford assay (Section 2.2.2).

6.2.3 Cultivation of Bacteria

Two strains of bacteria were used in this work. The pathogenic *Escherichia coli* O111:NM (ATCC 43887) and the probiotic *Lactobacillus rhamnosus* (ATCC 53103) were obtained from the American Type Culture Collection (ATCC), Manassas, VA, USA. All work was done aseptically under Bunsen flame. Liquid and agar media were sterilised by autoclave before use. Plate cultures were prepared by taking 50 uL of stock solution and streaking onto

plates. Plate cultures of enterohemorhaggic *Escherichia coli* O111:NM (ATCC 43887) were grown on Luria agar for 16 hours at 37 °C. Liquid cultures were prepared by transferring a single colony from plate cultures to 5 mL of Luria broth. Liquid cultures were placed on a shaker incubator at 37 °C, 250 rpm overnight.

Plate cultures of *Lactobacillus rhamnosus* (ATCC 53103) were grown on tryptone soya bean broth agar for three days at 37 °C in a candle jar to create semi-anaerobic conditions. Liquid cultures were prepared by transferring 2-3 colonies from culture plates to 5 mL of TSB media. Liquid cultures were placed on a shaker incubator at 37 °C, 250 rpm overnight. Liquid cultures of both bacterial strains were harvested by centrifugation at 4500 g for 5 mins and pellets were washed twice with PBS.

6.2.4 PVDF Plate Assay for Measuring Bacterial Adhesion to Intestinal Cell Membrane Proteins

Application of LS174T membrane proteins to the PVDF plate

The bacterial binding assay was performed on AcroWell™ 96-Well Membrane-Bottom Plates with BioTrace™ PVDF membrane (PALL corporation). The wells were wet with 100 µL of methanol then rinsed three times with 100 µL of PBS. LS174T membrane proteins (Section 6.2.2) were diluted to a concentration of 1 mg/mL and 50 µL was applied to each well. The plate was covered with a lid and allowed to sit for 30 mins to allow the membrane proteins to bind to the PVDF membrane on the bottom of the wells.

Preparation of bacterial dilutions and pre-incubation with human milk N-linked glycans

Bacteria (Section 6.2.3) were fluorescently stained by resuspending the cell pellets in 1mL of Sybr Green nucleic acid stain (prepared by mixing 1 mL of PBS with 1 µL of Sybr Green, Life Technologies). The bacteria were incubated on a shaker at 300 rpm for 3 mins covered in foil to minimise exposure to light. The cells were then centrifuged gently at 1000 g for 3 mins

and the supernatant removed. Cells were washed three times with 1 mL of PBS and pelleted by centrifugation at 1000 g for 3 mins each time. Bacteria were diluted to an OD600 of 2.0 with PBS using a biophotometer.

Preparation of N-linked glycan fractions: Human milk *N*-linked glycans were released from 10 mg of human milk proteins using the method developed in Chapter 3 (Section 3.2.6) and reduced using the one-pot method described in Chapter 4 (Section 4.2.1). Reduced glycans were purified by SPE on graphitised carbon and three fractions of *N*-linked glycans were obtained by sequential elution (Section 4.2.1). The "neutral" glycan fraction was eluted with 20% (v/v) acetonitrile. The "acidic" glycan fraction was then eluted with 40% (v/v) TFA acetonitrile, 0.05% (v/v). The "total" fraction was eluted with 40% (v/v) acetonitrile, 0.05% (v/v) TFA.

Pre-incubation of bacteria with N-linked glycans: The glycan fractions were prepared by diluting the glycans released from 0.5, 1 and 2 mg of human milk proteins in 75 μL aliquots of PBS then added to 75 μL of SYBR Green stained *E. coli* or *L. rhamnosus*. The mixtures were pre-incubated at room temperature for 15 mins on a shaker (~450 rpm) to minimise exposure to light. Controls were prepared by pre-incubating the bacteria in 75 μL of PBS only. All samples were prepared and performed in triplicate.

Application of bacterial strains to LS174T membrane proteins

The liquid containing excess LS174T membrane proteins was removed from the wells of the 96 well plates by pipette. The proteins bound to the PVDF wells were washed twice with 100 μL of PBS. The pre-incubated bacteria were then applied to the bound LS174T membrane proteins in the wells and incubated on a shaker (~250 rpm) at room temperature for 20 mins covered in foil to minimise exposure to light. Wells containing only membrane proteins and PBS (without bacteria) were also included to allow a background measurement of the fluorescence emitted by membrane alone. After incubation, the PBS and any unbound

bacteria were removed by pipette and wells were washed three times with 100 μL of PBS allowing 1 min on a plate shaker (~250 rpm) in between each wash step. Once all PBS was completely removed from the wells, bacterial binding to the membrane proteins was measured with a fluorescent plate reader (Fluostar, BMG technologies, Excitation at 485 nm, Emission at 520 nm). All samples were tested in triplicate. In addition, since it was found that *E. coli* or *L. rhamnosus* were stained to different extents by SYBR Green, 50 μL aliquots of the stained bacteria in PBS (OD600 of 1.0) were included in some additional wells on the plate to allow normalisation of the binding data according to the uptake of the stain.

6.3 Results and Discussion

6.3.1 Bacterial Adhesion to Human Intestinal Membrane Proteins following Pre-incubation with *N*-linked Glycans Released from Human Milk Proteins

A significant difference was observed in the initial adhesion of *E. coli* and *L. rhamnosus* to the human intestinal membrane proteins. After subtracting the background fluorescence of the membrane proteins and normalising to account for differences in the uptake of the fluorescent stain by the bacteria, *L. rhamnosus* was found to adhere more to the membrane proteins than the *E. coli* (p = 0.015) (Figure 6.1). This was in line with previous work in the laboratory and may be a reflection of the commensal relationship that *L. rhamnosus* has with humans; it seems that evolution has favoured the binding of healthy gut bacteria to the gut epithelial cells over that of the pathogenic competitors.

The adhesion of *E. coli* to the LS174T membrane proteins was actually increased following pre-incubation of the bacteria with the total, neutral and acidic *N*-linked glycans from human milk proteins (Figure 6.1). The Student's *t* test was used to determine the significance of the differences in adhesion, which were considered significant if $p < 0.05$. Pre-incubation with concentrations of *N*-linked glycans equivalent to those found in 0.5 mg milk protein did not significantly affect the binding of *E. coli* to the membrane proteins. However, pre-incubation in *N*-linked glycan-equivalent to 1 mg milk protein resulted in a significant increase in adhesion of *E. coli* for the total ($p = 0.02$), neutral ($p = 0.02$), and acidic ($p = 0.01$) fractions (Figure 6.1). Pre-incubation with *N*-linked glycan-equivalent to 2 mg milk protein resulted in a slight (but statistically insignificant) drop in adhesion in comparison to 1 mg milk protein, but still resulted in a significant increase in binding in comparison to binding of *E. coli* that was pre-incubated in PBS only (control) for the total ($p = 0.01$) and acidic ($p = 0.04$) fractions (Figure 6.1).
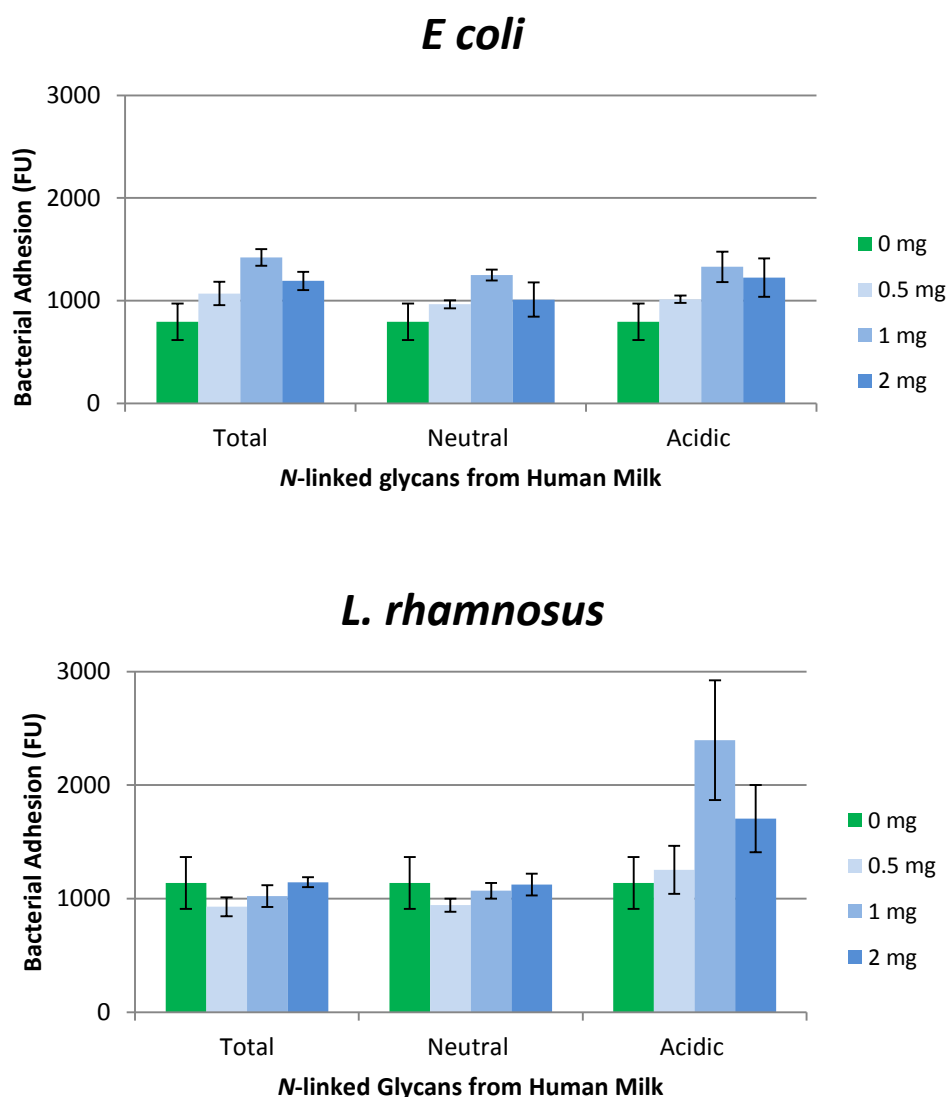
E coli

L. rhamnosus

Figure 6.1 – Adhesion of *E. coli* and *L. rhamnosus* to LS174T membrane proteins following pre-incubation with total, neutral and acidic *N*-linked glycans equivalent to those found in 0.5, 1 and 2 mg of human milk proteins. Adhesion is expressed in fluorescence units (FU). Error bars indicate one standard deviation above and below the mean based on three technical replicates.

Previous work in the laboratory using the same assay has found that the adhesion of *E. coli* 0111:NM to LS174T intestinal membrane proteins is inhibited by pre-incubation with human lactoferrin, the most abundant *N*-linked glycoprotein in human milk. Glycan involvement in the binding of human lactoferrin to a variety of human pathogens has also been demonstrated by other researchers via enzymatic release of terminal residues of the conjugated glycans altering the adhesion of the bacteria (Table 6.1). Therefore, the increased adhesion of *E. coli* to LS174T membrane proteins following pre-incubation with the *N*-linked glycans released from human milk in this work was surprising and suggests that the ability of the milk

glycoproteins to act as decoys (inhibitors) for adhesion of *E. coli* may not simply require the presence of the *N*-linked glycans alone; configuration of the *N*-linked glycans on the protein backbone may be integral to the interaction; pathogen binding to the protein backbone itself may also be involved.

This finding was consistent with work by Barboza *et al.* (2012), which reported a slight increase in the adhesion of motile (i.e. flagellated) *E. coli* O157:H7 to human Caco-2 colonic intestinal epithelial cells compared to the control following pre-incubation with the *N*-linked glycans released from human lactoferrin. In conjunction with the work reported by Barboza *et al.* (2012), the results obtained in this work suggest that the observation of the capacity of human milk lactoferrin to inhibit adhesion of *E. coli* strains to gastrointestinal cells or membrane proteins is more complex than simply an effect of either the *N*-linked glycans or the protein backbone alone.

In contrast, pre-incubation of *L. rhamnosus* with the total and neutral *N*-linked glycan fractions did not result in any significant changes in adhesion to intestinal membrane proteins (Figure 6.1). Pre-incubation of *L. rhamnosus* with the acidic *N*-linked glycan-equivalent to 1 mg (p = 0.02) and 2 mg (p = 0.04) milk protein also resulted in a significant increase in adhesion of the commensal bacterium to membrane proteins. Adhesion was greatest when *L. rhamnosus* was pre-incubated in the acidic *N*-linked glycan-equivalent to 1 mg milk protein and increasing the amount of glycans to that found in 2 mg milk protein resulted in a slight drop in adhesion in comparison.

Previous work in the laboratory has consistently shown that pre-incubation of *L. rhamnosus* in any glycan-containing compound including free *N*-linked glycans and glycoproteins results in increased adhesion of the commensal bacterium to human intestinal membrane proteins. The reason for this phenomenon is not fully understood, but appears compatible with a prebiotic mechanism in which the glycans promote the establishment of commensal bacteria in the gut

in favour of disease causing pathogens. The preparative scale quantities of *N*-linked glycans obtained from milk proteins in this work, and their separation into neutral and acidic fractions, enabled more detailed characterisation of the previously observed phenomenon. In particular, the work reported here (Figure 6.1) demonstrated that the released acidic *N*-linked glycans alone promote adhesion of *L. rhamnosus* to the membrane proteins, whilst neutral glycans did not. However, when acidic glycans were combined with neutral glycans (total glycan fraction), the pro-adhesive effect disappeared. This effect is difficult to explain at this stage but adds to the laboratory data in the frontier area of understanding the role of glycans in bacterial-host interactions and will be investigated further in future work, with the benefit of now having a protocol to obtain large enough quantities of fractions and individual structure to carry out the work.

6.4 Chapter Summary

The *N*-linked glycans released on a preparative scale from human milk proteins using the method developed in Chapter 3, and reduced using the method developed in Chapter 4, were applied to an established bacterial binding assay developed in-house to investigate their ability to affect the adhesion of bacteria to human intestinal membrane proteins. In this work, the *N*-linked glycans released from human milk proteins when used alone did not inhibit bacterial adhesion, suggesting that the anchoring of the *N*-linked glycans onto the protein backbone, as well as the protein backbone itself, may be critical to the anti-adhesive properties of human milk glycoproteins.

Pre-incubation of *L. rhamnosus* with the acidic *N*-linked glycans released from human milk proteins resulted in a significant increase in adhesion of *L. rhamnosus* to the intestinal membranes proteins. This finding suggested a potential mechanism by which the acidic *N*-linked glycans of human milk proteins promotes adhesion of *L. rhamnosus* to gut epithelial cells, by a mechanism which is not yet understood. The work in this chapter demonstrated the suitability of *N*-linked glycans released from human milk proteins using the method developed in Chapter 3, for application in *in vitro* biological assays.

Chapter 7 – Summary and Future Directions

7.1 Summary

In this work, human milk was used as a source of naturally produced protein *N*-linked glycans for potential use as standards to facilitate glycan analysis. To obtain the *N*-linked glycans from human milk, a method was developed to release the *N*-linked glycans on a preparative scale. It is well known that scale-up of a procedure involves more than simple multiplication. Switching to an in-solution release introduced unique challenges in order to maximise recovery of released glycans and maintain reproducibility. Scale up from the analytical scale involved changes at all levels of sample preparation. It was found that enzymatic release with PNGase F in solution and recovery by organic precipitation with acetone and subsequent extraction of the glycans with methanol could be used to obtain preparative scale amounts of *N*-linked glycans from milligram quantities of human milk proteins with little selective loss of glycan in comparison to the established analytical scale release. While the exact amount of *N*-linked glycans released could not be determined since destructive quantitation methods were avoided, a 1000-fold increase in the amount of starting material used was achieved, and the preparative scale release was shown to provide enough material for at least 100 analyses, in comparison to 1 analysis for the analytical scale release. The procedure developed in this section allowed for the release of N-linked glycans from mg quantities of proteins, which differs from currents methods that are designed for μg quantities.

In order to obtain pure glycan structures, a preparative scale technique using PGC-HPLC was developed to separate the neutral *N*-linked glycans into unique structures. An offline detection of the eluted glycans was established using MALDI-TOF MS that resulted in the collection of six neutral *N*-linked glycans at approximately 90% purity.

Structural characterisation of the six purified neutral *N*-linked glycans by analytical PGC-LC-ESI-MS fragmentation identified a set of fucosylated *N*-linked glycans that are not

commercially available. The six purified *N*-linked glycans were used to establish a retention time library of fucosylated structures and it was demonstrated that an α1 − 6 linked core fucose resulted in an increase in retention on PGC, while the terminal Lewis x epitope (Galβ1 − 4[Fucα1 − 3]GlcNAc) resulted in a decrease in retention on PGC. The availability of purified structures in sufficient quantities enabled these types of studies to be carried out.

The availability of large scale quantities of *N*-linked glycans released from human milk allows their use in other downstream applications such as biological assays. An established bacterial binding assay developed in-house was used to investigate the potential of the *N*-linked glycans released from human milk proteins to influence the binding of gut bacteria to human intestinal cells. In this study, it was demonstrated that *N*-linked glycans released from human milk proteins did not reduce the adhesion of *E. coli* to human intestinal epithelial cell membrane proteins suggesting the requirement of both glycan and protein components in the anti-adhesive activity attributed to the human milk glycoproteins. It was also demonstrated that the acidic *N*-linked glycans released from human milk proteins caused an increase in the adhesion of *L. rhamnosus* to human intestinal epithelial cell membrane proteins.

In conclusion, a set of techniques were developed in this work to enable the preparative scale release of *N*-linked glycans from human milk glycoproteins, and separation of the released *N*-linked glycans into purified well characterised structures for potential use as standards. The techniques developed were designed for use in a laboratory setting and potentially allow the general production of *N*-linked glycan standards from natural sources.

7.1 Future Directions

The preparative scale release protocol developed in this work was specifically designed for human milk. Further optimisation of the protocol could potentially allow the application of the preparative scale release to purified glycoproteins or other human secretions such as saliva. Additionally, accurate quantitation needs to be performed in order to validate the

efficiency and yield of the preparative scale release. Potential quantitation techniques include fluorescent and isotopic labelling techniques. Fluorescent labelling involves the derivatisation of the reducing end of glycans with chemical labels that contain a fluorophore such as 2-AB (Ruhaak et al., 2010). By including a labelled internal standard of a known amount, relative quantitation can be achieved by LC separation coupled with fluorescent detection. However, since the purified *N*-linked glycans were reduced before separation in this work, they lacked the potential for quantitation by fluorescent labelling.

Isotopic labelling involves the derivatisation of glycans with stable isotopes such as $^{12}C_6$ aniline to the reducing end, or ^{13}C methyl iodide for permethylation, or sodium borodeuteride ($NaBD_4$) for reduction, which can then be analysed by MS to resolve isotopic mass differences (Xia et al., 2009). Quantitation is thus achieved by comparing samples derivatised with the "heavy" label with samples derivatised with the "light" label (Alvarez-Manilla et al., 2007). As an alternative to reducing end labels, permethylation with a heavy isotope provides a potential method for the quantitation of reduced *N*-linked glycans, such as those obtained by the preparative scale technique developed in this work.

While six neutral *N*-linked glycans were purified by PGC-HPLC, a number of mixed fractions were also collected and could be further separated to obtain a more complete set of human milk protein glycan standards. Furthermore, the optimisation of a suitable MALDI-TOF detection method for the acidic *N*-linked glycans could also facilitate the preparative scale separation and collection of unique acidic *N*-linked glycan structures using the PGC-HPLC technique developed in this work.

The capacity to separate neutral from acidic fractions of human milk glycans by PGC demonstrated that the *N*-linked glycans released from human milk proteins did not reduce the adhesion of *E. coli* to human intestinal epithelial cell membrane proteins as has been shown by milk glycoproteins but further investigation of the phenomenon was beyond the scope of

this thesis. There are also many other gastrointestinal pathogens that need to studied in terms of the glycan-binding mechanisms. These investigations could be performed using the PVDF binding assay described in this work, or with other methods such as flow cytometry, fluorescent microscopy or carbohydrate microarrays. The preparative scale release developed in this work also provides a useful method for obtaining large quantities of *N*-linked glycans, suitable for these *in vitro* applications.

References

Alpert, A.J., M. Shukla, A.K. Shukla, L.R. Zieske, S.W. Yuen, M.A. Ferguson, A. Mehlert, M. Pauly, and R. Orlando. 1994. Hydrophilic-interaction chromatography of complex carbohydrates. *Journal of Chromatography A*. 676:191-122.

Alvarez-Manilla, G., N.L. Warren, T. Abney, J. Atwood, P. Azadi, W.S. York, M. Pierce, and R. Orlando. 2007. Tools for glycomics: relative quantitation of glycans by isotopic permethylation using 13CH3I. *Glycobiology*. 17:677-687.

Apweiler, R., H. Hermjakob, and N. Sharon. 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 1473:4-8.

Arnold, J.N., M.R. Wormald, R.B. Sim, P.M. Rudd, and R.A. Dwek. 2007. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annual Review of Immunology*. 25:21-50.

Baldwin, T.H., L.R. P., J.M.J. Milda, O. Ryoichi, K.I. Celia, and K.D. Barry. 1976. Human colonic adenocarcinoma cells. *In Vitro*. 12:180-191.

Barboza, M., J. Pinzon, S. Wickramasinghe, J.W. Froehlich, I. Moeller, J.T. Smilowitz, L.R. Ruhaak, J. Huang, B. Lönnerdal, J.B. German, J.F. Medrano, B.C. Weimer, and C.B. Lebrilla. 2012. Glycosylation of human milk lactoferrin exhibits dynamic changes during early lactation enhancing its role in pathogenic bacteria-host interactions. *Molecular & Cellular Proteomics*. 11.

Bennion, B.J., and V. Daggett. 2003. The molecular basis for the chemical denaturation of proteins by urea. *Proceedings of the National Academy of Sciences*. 100:5142-5147.

Bode, L. 2012. Human milk oligosaccharides: Every baby needs a sugar mama. *Glycobiology*. 22:1147-1162.

Boscher, C., J.W. Dennis, and I.R. Nabi. 2011. Glycosylation, galectins and cellular signaling. *Current Opinion in Cell Biology*. 23:383-392.

Brockhausen, I. 1999. Pathways of O-glycan biosynthesis in cancer cells. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 1473:67-95.

Brownlee, M.D. 1995. Advanced protein glycosylation in daibetes and aging *Annual Review of Medicine*. 46:223-234.

Ceroni, A., A. Dell, and S.M. Haslam. 2007. The GlycanBuilder: a fast, intuitive and flexible software tool for building and displaying glycan structures. *Source code for biology and medicine*. 2:3.

Ceroni, A., K. Maass, H. Geyer, R. Geyer, A. Dell, and S.M. Haslam. 2008. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *Journal of Proteome Research*. 7:1650-1659.

Chaturvedi, P., C.D. Warren, M. Altaye, A.L. Morrow, G. Ruiz-Palacios, L.K. Pickering, and D.S. Newburg. 2001. Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation. *Glycobiology*. 11:365-372.

Choi, S.-S., and S.-H. Ha. 2006. Characterization of Ionized Maltooligosaccharides by Sodium Cation in MALDI-TOFMS Depending on the Molecular Size. *Bulletin of the Korean Chemical Society*. 27:1243-1245.

Colley, K.J. 1997. Golgi localization of glycosyltransferases: more questions than answers. *Glycobiology*. 7:1-13.

Cooper, C.A., E. Gasteiger, and N.H. Packer. 2001. GlycoMod – A software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics*. 1:340-349.

Coppa, G.V., O. Gabrielli, P. Pierani, C. Catassi, A. Carlucci, and P.L. Giorgi. 1993. Changes in Carbohydrate Composition in Human Milk Over 4 Months of Lactation. *Pediatrics*. 91:637-641.

Cravioto, A., A. Tello, H. Villafan, J. Ruiz, S. del Vedovo, and J.R. Neeser. 1991. Inhibition of localized adhesion of enteropathogenic *Escherichia coli* to HEp-2 cells by immunoglobulin and oligosaccharide fractions of human colostrum and breast milk. *The Journal of Infectious Diseases*. 163:1247.

Dallas, D.C., W.F. Martin, J.S. Strum, A.M. Zivkovic, J.T. Smilowitz, M.A. Underwood, M. Affolter, C.B. Lebrilla, and J.B. German. 2011. N-Linked Glycan Profiling of Mature Human Milk by High-Performance Microfluidic Chip Liquid Chromatography Time-of-Flight Tandem Mass Spectrometry. *Journal of Agricultural and Food Chemistry*. 59:4255-4263.

de Vries, T., R.M.A. Knegtel, E.H. Holmes, and B.A. Macher. 2001. Fucosyltransferases: structure/function studies. *Glycobiology*. 11:119-128.

Domon, B., and C. Costello. 1988. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J*. 5:397-409.

Dwek, R.A. 1996. Glycobiology: toward understanding the function of sugars. *Chemical Reviews*. 96:683-720.

Dziuba, J., and P. Minkiewicz. 1996. Influence of glycosylation on micelle-stabilizing ability and biological properties of C-terminal fragments of cow's κ-casein. *International Dairy Journal*. 6:1017-1044.

Echols, K.R., R.W. Gale, K. Feltz, J. O'Laughlin, D.E. Tillitt, and T.R. Schwartz. 1998. Loading capacity and chromatographic behavior of a porous graphitic carbon column for polychlorinated biphenyls. *Journal of Chromatography A*. 811:135-144.

El-Rassi, Z. 2002. Carbohydrate Analysis by Modern Chromatography and Electrophoresis. Elsevier Science.

Everest-Dass, A.V., J.L. Abrahams, D. Kolarich, N.H. Packer, and M.P. Campbell. 2013. Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *Journal of the American Society for Mass Spectrometry*. 24:895-906.

Evers, D.L., R.L. Hung, V.H. Thomas, and K.G. Rice. 1998. Preparative Purification of a High-Mannose TypeN-Glycan from Soy Bean Agglutinin by Hydrazinolysis and Tyrosinamide Derivatization. *Analytical Biochemistry*. 265:313-316.

Falk, P., K.A. Roth, T. Boren, T.U. Westblom, J.I. Gordon, and S. Normark. 1993. An in vitro adherence assay reveals that *Helicobacter pylori* exhibits cell lineage-specific tropism in the human gastric epithelium. *Proceedings of the National Academy of Sciences of the United States of America*. 90:2035-2039.

Fountoulakis, M. 2001. Proteomics: Current technologies and applications in neurological disorders and toxicology. *Amino Acids*. 21:363-381.

Fox, P.F., and A. Brodkorb. 2008. The casein micelle: historical aspects, current concepts and significance. *International Dairy Journal*. 18:677-684.

Froehlich, J.W., E.D. Dodds, M. Barboza, E.L. McJimpsey, R.R. Seipert, J. Francis, H.J. An, S. Freeman, J.B. German, and C.B. Lebrilla. 2010. Glycoprotein expression in human milk during lactation. *Journal of Agricultural and Food Chemistry*. 58:6440-6448.

Gilbert, M.T., and J.H. Knox. 1981. Preparation of porous carbon. US Patent US4263268A.

Ginger, M.R., and M.R. Grigor. 1999. Comparative aspects of milk caseins. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*. 124:133-145.

Gritti, F., and G. Guiochon. 2012. Overload behavior and apparent efficiencies in chromatography. *Journal of Chromatography A*. 1254:30-42.

Grootaert, C., N. Boon, F. Zeka, B. Vanhoecke, M. Bracke, W. Verstraete, and T. Van de Wiele. 2011. Adherence and viability of intestinal bacteria to differentiated Caco-2 cells quantified by flow cytometry. *Journal of Microbiological Methods*. 86:33-41.

Grulee, C.G., H.N. Sanford, and H. Schwartz. 1935. Breast and artificially fed infants: A study of the age incidence in the morbidity and mortality in twenty thousand cases. *Journal of the American Medical Association*. 104:1986-1988.

Günther, H. 2013. NMR Spectroscopy: Basic Principles, Concepts and Applications in Chemistry. Wiley.

Guo, C., B.E. Campbell, K. Chen, A.M. Lenhoff, and O.D. Velev. 2003. Casein precipitation equilibria in the presence of calcium ions and phosphates. *Colloids and Surfaces B: Biointerfaces*. 29:297-307.

Hager, D.A., and R.R. Burgess. 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with sigma subunit of Escherichia coli RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. *Analytical Biochemistry*. 109:76-86.

Hakomori, S.-i. 1996. Tumor Malignancy Defined by Aberrant Glycosylation and Sphingo(glyco)lipid Metabolism. *Cancer Research*. 56:5309-5318.

Han, N.S., T.-J. Kim, Y.-C. Park, J. Kim, and J.-H. Seo. 2012. Biotechnological production of human milk oligosaccharides. *Biotechnology Advances*. 30:1268-1278.

Hao, P., Y. Ren, and Y. Xie. 2010. An improved protocol for N-glycosylation analysis of gel-separated sialylated glycoproteins by MALDI-TOF/TOF. *Public Library of Science One*. 5:e15096.

Harvey, D. 2005. Fragmentation of negative ions from carbohydrates: Part 3. Fragmentation of hybrid and complex N-linked glycans. *Journal of the American Society for Mass Spectrometry*. 16:647-659.

Harvey, D.J. 2000. Electrospray mass spectrometry and fragmentation of N-linked carbohydrates derivatized at the reducing terminus. *Journal of the American Society for Mass Spectrometry*. 11:900-915.

Harvey, D.J. 2003. Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *International Journal of Mass Spectrometry*. 226:1-35.

Harvey, D.J., A.H. Merry, L. Royle, M. P. Campbell, R.A. Dwek, and P.M. Rudd. 2009. Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *Proteomics*. 9:3796-3801.

Harvey, D.J., L. Royle, C.M. Radcliffe, P.M. Rudd, and R.A. Dwek. 2008. Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. *Analytical Biochemistry*. 376:44-60.

Herbert, C.G., and R.A.W. Johnstone. 2010. Mass Spectrometry Basics. Taylor & Francis.

Hubbard, S.C., and R.J. Ivatt. 1981. Synthesis and processing of asparagine-linked oligosaccharides. *Annual Review of Biochemistry*. 50:555-583.

Jensen, P.H., N.G. Karlsson, D. Kolarich, and N.H. Packer. 2012. Structural analysis of N- and O-glycans released from glycoproteins. *Nature Protocols*. 7:1299-1310.

Jensen, R.G. 1995. Handbook of Milk Composition. Elsevier Science.

Jensen, R.G. 1999. Lipids in human milk. *Lipids*. 34:1243-1271.

Jiang, L., L. He, and M. Fountoulakis. 2004. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *Journal of Chromatography A*. 1023:317-320.

Jung, S.-Y., Y. Liu, and C.P. Collier. 2008. Fast Mixing and Reaction Initiation Control of Single-Enzyme Kinetics in Confined Volumes. *Langmuir*. 24:4439-4442.

Kandzia, S., and J. Costa. 2013. N-Glycosylation Analysis by HPAEC-PAD and Mass Spectrometry. *In* Ovarian Cancer. Vol. 1049. A. Malek and O. Tchernitsa, editors. Humana Press. 301-312.

Karlsson, K.A. 1995. Microbial recognition of target-cell glycoconjugates. *Current Opinion in Structural Biology*. 5:622-635.

Kelley, P.E., J.N. Louris, W.E. Reynolds, G.C. Stafford, and J.E.P. Syka. 1988. Method of operating ion trap detector in MS/MS mode. US Patent US4736101A.

Kelly, C.G., and J.S. Younson. 2000. Anti-adhesive strategies in the prevention of infectious disease at mucosal surfaces. *Expert Opinion on Investigational Drugs*. 9:1711-1721.

Kornfeld, R., and S. Kornfeld. 1985. Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry*. 54:631-664.

Kunz, C., and B. Lönnerdal. 1990. Human-milk proteins: analysis of casein and casein subunits by anion-exchange chromatography, gel electrophoresis, and specific staining methods. *The American Journal of Clinical Nutrition*. 51:37-46.

Lægreid, A., and J. Fuglesang. 1986. Human and bovine milk: comparison of ganglioside composition and enterotoxin-inhibitory activity. *Pediatric Research*. 20:416-421.

Lauc, G., and V. Zoldos. 2010. Protein glycosylation-an evolutionary crossroad between genes and environment. *Molecular BioSystems*. 6:2373-2379.

le Maire, M., P. Champeil, and J.V. Møller. 2000. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 1508:86-111.

Le Pendu, J. 2004. Histo-blood group antigen and human milk oligosaccharides: genetic polymorphism and risk of infectious diseases. *Advances in Experimental Medicine and Biology*. 554:135-143.

Leblond, C.P., R.E. Glegg, and D. Eidinger. 1957. Presence of carbohydrates with free 1,2-glycol groups in sites stained by period acid-schiff techcnique *Journal of Histochemistry & Cytochemistry*. 5:445-458.

Lepenies, B., J. Yin, and P.H. Seeberger. 2010. Applications of synthetic carbohydrates to chemical biology. *Current Opinion in Chemical Biology*. 14:404-411.

Lesman-Movshovich, E., B. Lerrer, and N. Gilboa-Garber. 2003. Blocking of *Pseudomonas aeruginosa* lectins by human milk glycans. *Canadian Journal of Microbiology*. 49:230-235.

Leymarie, N., and J. Zaia. 2012. Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Analytical Chemistry*. 84:3040-3048.

Liao, Y., R. Alvarado, B. Phinney, and B. Lönnerdal. 2011. Proteomic characterization of human milk whey proteins during a twelve-month lactation period. *Journal of Proteome Research*. 10:1746-1754.

Liu, B., and D.S. Newburg. 2013. Human milk glycoproteins protect infants against human pathogens. *Breastfeeding medicine : the official journal of the Academy of Breastfeeding Medicine*.

Liu, Y., and R. Guo. 2008. pH-dependent structures and properties of casein micelles. *Biophysical Chemistry*. 136:67-73.

Lönnerdal, B. 2003. Nutritional and physiologic significance of human milk proteins. *The American Journal of Clinical Nutrition*. 77:1537S-1543S.

Lough, W.J., and I.W. Wainer. 1995. High Performance Liquid Chromatography: Fundamental Principles and Practice. Taylor & Francis.

Macias, C., and F.J. Schweigert. 2001. Changes in the concentration of carotenoids, vitamin A, alpha-tocopherol and total lipids in human milk throughout early lactation. *Annals of Nutrition and Metabolism*. 45:82-85.

Madadlou, A., M.E. Mousavi, Z. Emam-Djomeh, D. Sheehan, and M. Ehsani. 2009. Alkaline pH does not disrupt re-assembled casein micelles. *Food Chemistry*. 116:929-932.

Maley, F., R.B. Trimble, A.L. Tarentino, and T.H. Plummer Jr. 1989. Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Analytical Biochemistry*. 180:195-204.

Marino, K., J. Bones, J.J. Kattla, and P.M. Rudd. 2010. A systematic approach to protein glycosylation analysis: a path through the maze. *Nature Chemical Biology*. 6:713-723.

Melmer, M., T. Stangler, A. Premstaller, and W. Lindner. 2011. Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis. *Journal of Chromatography A*. 1218:118-123.

Morelle, W., V. Faid, F. Chirat, and J.-C. Michalski. 2009. Analysis of N- and O-Linked glycans from glycoproteins using MALDI-TOF mass spectrometry. *In* Glycomics. Vol. 534. N. Packer and N. Karlsson, editors. Humana Press. 3-21.

Morita, H., H. Toh, K. Oshima, M. Murakami, T.D. Taylor, S. Igimi, and M. Hattori. 2009. Complete genome sequence of the probiotic *Lactobacillus rhamnosus* ATCC 53103. *Journal of Bacteriology*. 191:7630-7631.

Morrow, A.L., and J.M. Rangel. 2004. Human milk protection against infectious diarrhea: implications for prevention and clinical care. *Seminars in Pediatric Infectious Diseases*. 15:221-228.

Naarding, M.A., A.M. Dirac, I.S. Ludwig, D. Speijer, S. Lindquist, E.L. Vestman, M.J. Stax, T.B. Geijtenbeek, G. Pollakis, O. Hernell, and W.A. Paxton. 2006. Bile salt-stimulated lipase from human milk binds DC-SIGN and inhibits human immunodeficiency virus type 1 transfer to CD4+ T cells. *Antimicrob Agents and Chemotherapy*. 50:3367-3374.

Newburg, D.S. 2009. Neonatal protection by an innate immune system of human milk consisting of oligosaccharides and glycans. *Journal of Animal Science*. 87:26-34.

Newburg, D.S., L.K. Pickering, R.H. McCluer, and T.G. Cleary. 1990. Fucosylated oligosaccharides of human milk protect suckling mice from heat-stabile enterotoxin of *Escherichia coli*. *The Journal of Infectious Diseases*. 162:1075-1080.

Newburg, D.S., G.M. Ruiz-Palacios, M. Altaye, P. Chaturvedi, J. Meinzen-Derr, M.d.L. Guerrero, and A.L. Morrow. 2004. Innate protection conferred by fucosylated oligosaccharides of human milk against diarrhea in breastfed infants. *Glycobiology*. 14:253-263.

Newburg, D.S., G.M. Ruiz-Palacios, and A.L. Morrow. 2005. Human milk glycans protect infants against enteric pathogens *Annual Review of Nutrition*. 25:37-58.

Nizet, V., A. Smith, P. Sullam, and C. Rubens. 1998. A simple microtiter plate screening assay for bacterial invasion or adherence. *Methods Cell Sci*. 20:107-111.

Nwosu, C.C., D.L. Aldredge, H. Lee, L.A. Lerno, A.M. Zivkovic, J.B. German, and C.B. Lebrilla. 2012. Comparison of the human and bovine milk N-glycome via high-performance microfluidic chip liquid chromatography and tandem mass spectrometry. *Journal of Proteome Research*. 11:2912-2924.

O'Donnell, R., J.W. Holland, H.C. Deeth, and P. Alewood. 2004. Milk proteomics. *International Dairy Journal*. 14:1013-1023.

Orlando, R. 2010. Quantitative Glycomics. *In* Functional Glycomics. Vol. 600. J. Li, editor. Humana Press. 31-49.

Pabst, M., J.S. Bondili, J. Stadlmann, L. Mach, and F. Altmann. 2007. Mass + Rrtention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. *Analytical Chemistry*. 79:5051-5057.

Pabst, M., S.Q. Wu, J. Grass, A. Kolb, C. Chiari, H. Viernstein, F.M. Unger, F. Altmann, and S. Toegel. 2010. IL-1beta and TNF-alpha alter the glycophenotype of primary human chondrocytes in vitro. *Carbohydr Research*. 345:1389-1393.

Packer, N.H., M.A. Lawson, D.R. Jardine, and J.W. Redmond. 1998. A general approach to desalting oligosaccharides released from glycoproteins. *Glycoconjugate J*. 15:737-747.

Packer, N.H., C.W. von der Lieth, K.F. Aoki-Kinoshita, C.B. Lebrilla, J.C. Paulson, R. Raman, P. Rudd, R. Sasisekharan, N. Taniguchi, and W.S. York. 2008. Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11-13, 2006). *Proteomics*. 8:8-20.

Papac, D.I., J.B. Briggs, E.T. Chin, and A.J.S. Jones. 1998. A high-throughput microscale method to release N-linked oligosaccharides from glycoproteins for matrix-assisted laser desorption/ionization time-of-flight mass spectrometric analysis. *Glycobiology*. 8:445-454.

Papac, D.I., A. Wong, and A.J.S. Jones. 1996. Analysis of acidic oligosaccharides and glycopeptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Analytical Chemistry*. 68:3215-3223.

Patenaude, S.I., N.O.L. Seto, S.N. Borisova, A. Szpacenko, S.L. Marcus, M.M. Palcic, and S.V. Evans. 2002. The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nature Structural Biology*. 9:685-690.

Pereira, L. 2008. Porous graphitic carbon as a stationary phase in HPLC: theory and applications. *Journal of Liquid Chromatography & Related Technologies*. 31:1687-1731.

Peterson, R., W.Y. Cheah, J. Grinyer, and N. Packer. 2013. Glycoconjugates in human milk: protecting infants from disease. *Glycobiology*. 23:1425-1438.

Picciano, M.F. 2001. Nutrient composition of human milk. *The Pediatric Clinics of North America*. 48:53-67.

Plummer, T.H., and A.L. Tarentino. 1981. Facile cleavage of complex oligosaccharides from glycopeptides by almond emulsin peptide: N-glycosidase. *Journal of Biological Chemistry*. 256:10243-10246.

Pratt, M.R., and C.R. Bertozzi. 2005. Synthetic glycopeptides and glycoproteins as tools for biology. *Chemical Society Reviews*. 34:58-68.

Rademacher, T.W., R.B. Parekh, and R.A. Dwek. 1988. Glycobiology. *Annual Review of Biochemistry*. 57:785-838.

Raiha, N.C. 1985. Nutritional proteins in milk and the protein requirement of normal infants. *Pediatrics*. 75:136-141.

Rakus, J.F., and L.K. Mahal. 2011. New technologies for glycomic analysis: toward a systematic understanding of the glycome. *Annual Review of Analytical Chemistry*. 4:367-392.

Ramsby, M.L., G.S. Makowski, and E.A. Khairallah. 1994. Differential detergent fractionation of isolated hepatocytes: Biochemical, immunochemical and two-dimensional gel electrophoresis characterization of cytoskeletal and noncytoskeletal compartments. *Electrophoresis*. 15:265-277.

Rathore, A., and A. Velayudhan. 2002. Scale-Up and Optimization in Preparative Chromatography: Principles and Biopharmaceutical Applications. Taylor & Francis.

Reis, C.A., H. Osorio, L. Silva, C. Gomes, and L. David. 2010. Alterations in glycosylation as biomarkers for cancer detection. *Journal of Clinical Pathology*. 63:322-329.

Rolfe, R.D. 2000. The role of probiotic cultures in the control of gastrointestinal health. *The Journal of Nutrition*. 130:396.

Rolfe, R.D., and W. Song. 1995. Immunoglobulin and non-immunoglobulin components of human milk inhibit *Clostridium difficile* toxin A-receptor binding. *Journal of medical microbiology*. 42:10-19.

Royle, L., E. Matthews, A. Corfield, M. Berry, P. Rudd, R. Dwek, and S. Carrington. 2008. Glycan structures of ocular surface mucins in man, rabbit and dog display species differences. *Glycoconjugate J*. 25:763-773.

Royle, L., C. Radcliffe, R. Dwek, and P. Rudd. 2007. Detailed Structural Analysis of N-Glycans Released From Glycoproteins in SDS-PAGE Gel Bands Using HPLC Combined With

Exoglycosidase Array Digestions. *In* Glycobiology Protocols. Vol. 347. I. Brockhausen, editor. Humana Press.

Rudloff, S., and C. Kunz. 1997. Protein and nonprotein nitrogen components in human milk, bovine milk, and infant formula: quantitative and qualitative aspects in infant nutrition. *Journal of Pediatric Gastroenterology and Nutrition*. 24:328-344.

Ruhaak, L., A. Deelder, and M. Wuhrer. 2009. Oligosaccharide analysis by graphitized carbon liquid chromatography–mass spectrometry. *Analytical and Bioanalytical Chemistry*. 394:163-174.

Ruhaak, L., G. Zauner, C. Huhn, C. Bruggink, A. Deelder, and M. Wuhrer. 2010. Glycan labeling strategies and their use in identification and quantification. *Analytical and Bioanalytical Chemistry*. 397:3457-3481.

Saint, L., M. Smith, and P.E. Hartmann. 1984. The yield and nutrient content of colostrum and milk of women from giving birth to 1 month post-partum. *British Journal of Nutrition*. 52:87-95.

Samyn-Petit, B., J.-P. Wajda Dubos, F. Chirat, B. Coddeville, G. Demaizieres, S. Farrer, M.-C. Slomianny, M. Theisen, and P. Delannoy. 2003. Comparative analysis of the site-specific N-glycosylation of human lactoferrin produced in maize and tobacco plants. *European Journal of Biochemistry*. 270:3235-3242.

Scanlin, T.F., and M.C. Glick. 1999. Terminal glycosylation in cystic fibrosis. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 1455:241-253.

Schroten, H., R. Plogmann, F.G. Hanisch, J. Hacker, R. Nobis-Bosch, and V. Wahn. 1993. Inhibition of adhesion of S-fimbriated *E. coli* to buccal epithelial cells by human skim milk is predominantly mediated by mucins and depends on the period of lactation. *Acta Pædiatrica*. 82:6-11.

Sharon, N. 2006. Carbohydrates as future anti-adhesion drugs for infectious diseases. *Anglais*. 1760:527-537.

Shi, Y.-d., G.-q. Sun, Z.-g. Zhang, X. Deng, X.-h. Kang, Z.-d. Liu, Y. Ma, and Q.-h. Sheng. 2011. The chemical composition of human milk from Inner Mongolia of China. *Food Chemistry*. 127:1193-1198.

Smythe, C.V. 1936. The Reaction of Iodoacetate and of Iodoacetamide with various Sulfhydryl group, with Urease, and with Yest PreparationsTHE *Journal of Biological Chemistry*. 114:601-612.

Stadlmann, J., M. Pabst, D. Kolarich, R. Kunert, and F. Altmann. 2008. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics*. 8:2858-2871.

Stahl, B., S. Thurl, J.R. Zeng, M. Karas, F. Hillenkamp, M. Steup, and G. Sawatzki. 1994. Oligosaccharides from human milk as revealed by matrix-assisted laser desorption/ionization mass spectrometry. *Analytical Biochemistry*. 223:218-226.

Stronge, V.S., Y. Saito, Y. Ihara, and D.B. Williams. 2001. Relationship between calnexin and BiP in suppressing aggregation and promoting refolding of protein and glycoprotein substrates. *Journal of Biological Chemistry*. 276:39779-39787.

Stubbs, H.J., M.A. Shia, and K.G. Rice. 1997. Preparative Purification of Tetraantennary Oligosaccharides from Human Asialyl Orosomucoid. *Analytical Biochemistry*. 247:357-365.

Szabo, Z., A.s. Guttman, and B.L. Karger. 2010. Rapid release of N-linked glycans from glycoproteins by pressure-cycling technology. *Analytical Chemistry*. 82:2588-2593.

Takasaki, S., T. Mizuochi, and A. Kobata. 1982. Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Methods in Enzymology*. 83:263-268.

Tarentino, A.L., C.M. Gomez, and T.H. Plummer. 1985. Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Biochemistry*. 24:4665-4671.

Tretter, V., F. Altmann, and L. MÄRz. 1991. Peptide-N4-(N-acetyl-β-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached $\alpha 1 \rightarrow 3$ to the asparagine-linked N-acetylglucosamine residue. *European Journal of Biochemistry*. 199:647-652.

Unger, K.K. 1983. Porous carbon packings for liquid chromatography. *Analytical Chemistry* 55:361A-375A.

Van Klinken, B.J., A.W. Einerhand, H.A. Buller, and J. Dekker. 1998. Strategic biochemical analysis of mucins. *Analytical Biochemistry*. 265:103-116.

Varki, A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*. 3:97-130.

Varki, A. 2009. Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Verostek, M.F., C. Lubowski, and R.B. Trimble. 2000. Selective Organic Precipitation/Extraction of Released N-Glycans Following Large-Scale Enzymatic Deglycosylation of Glycoproteins. *Analytical Biochemistry*. 278:111-122.

Wakabayashi, H., K. Yamauchi, and M. Takase. 2006. Lactoferrin research, technology and applications. *International Dairy Journal*. 16:1241-1251.

Wang, J., P. Sporns, and N.H. Low. 1999. Analysis of food oligosaccharides using MALDI-MS: quantification of fructooligosaccharides. *Journal of Agricultural and Food Chemistry*. 47:1549-1557.

Waugh, D.F., and P.H. von Hippel. 1956. κ-casein and the stabilization of casein micelles. *Journal of the American Chemical Society*. 78:4576-4582.

Wessel, D., and U.I. Flügge. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Analytical Biochemistry*. 138:141-143.

Wilm, M. 2011. Principles of Electrospray Ionization. *Molecular & Cellular Proteomics*. 10:M111.009407.

Wojcik, K.Y., D.J. Rechtman, M.L. Lee, A. Montoya, and E.T. Medo. 2009. Macronutrient analysis of a nationwide sample of donor breast milk. *Journal of the American Dietetic Association*. 109:137-140.

Wold, A.E., J. Mestecky, M. Tomana, A. Kobata, H. Ohbayashi, T. Endo, and C.S. Eden. 1990. Secretory immunoglobulin A carries oligosaccharide receptors for *Escherichia coli* type 1 fimbrial lectin. *Infection and Immunity*. 58:3073-3077.

Wu, A.M., E.A. Kabat, B. Nilsson, D.A. Zopf, F.G. Gruezo, and J. Liao. 1984. Immunochemical studies on blood groups. Purification and characterization of radioactive 3H-reduced di- to hexasaccharides produced by alkaline beta-elimination-borohydride 3H reduction of Smith degraded blood group A active glycoproteins. *Journal of Biological Chemistry*. 259:7178-7186.

Wyskovsky, W. 1998. Enzymatic reactions in small spatial volumes: comment on a model of Hess and Mikhailov. *Biophysical Chemistry*. 71:73-81.

Xia, B., C.L. Feasley, G.P. Sachdev, D.F. Smith, and R.D. Cummings. 2009. Glycan reductive isotope labeling for quantitative glycomics. *Analytical Biochemistry*. 387:162-170.

Xia, X., J. Meng, P.F. McDermott, S. Ayers, K. Blickenstaff, T.T. Tran, J. Abbott, J. Zheng, and S. Zhao. 2010. Presence and characterization of shiga toxin-producing *Escherichia coli* and other potentially diarrheagenic *E. coli* strains in retail meats. *Applied and Environmental Microbiology*. 76:1709-1717.

Yamawaki, N., M. Yamada, T. Kan-no, T. Kojima, T. Kaneko, and A. Yonekubo. 2005. Macronutrient, mineral and trace element composition of breast milk from Japanese women. *Journal of Trace Elements in Medicine and Biology*. 19:171-181.

Zaia, J. 2008. Mass Spectrometry and the Emerging Field of Glycomics. *Chemistry & Biology*. 15:881-892.

Zaia, J. 2010. Mass spectrometry and glycomics. *Omics : a journal of integrative biology*. 14:401-418.

Zana, R. 1980. Ionization of cationic micelles: Effect of the detergent structure. *Journal of Colloid and Interface Science*. 78:330-337.

Zauner, G., A.M. Deelder, and M. Wuhrer. 2011. Recent advances in hydrophilic interaction liquid chromatography (HILIC) for structural glycomics. *Electrophoresis*. 32:3456-3466.

Zellner, M., W. Winkler, H. Hayden, M. Diestinger, M. Eliasen, B. Gesslbauer, I. Miller, M. Chang, A. Kungl, E. Roth, and R. Oehler. 2005. Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets. *Electrophoresis*. 26:2481-2489.

Appendix I

Protocol for the Preparative Scale Release of *N*-linked Glycans from Human Milk Proteins

Materials

1. 50 mL Falcon tube
2. Mammalian protease inhibitor cocktail (Novagen)
3. Centrifuge capable of 10 000 rcf, 4 °C (Sigma, 8-16K)
4. -20 °C Freezer
5. pH meter
6. 1M Hydrochloric acid (HCl; Ajax finechem 36%, 1367)
7. 1M Calcium chloride (CaCl$_2$; Sigma Aldrich, C3881)
8. Centrifuge capable of 100 000 g, 4 °C (Sorvall, Discovery ultracentrifuge 90SE)
9. Methanol (MeOH; Merck, Lichrosolve®,106018)
10. Chloroform (CHCl$_3$; Sigma Aldrich, 650498)
11. Deionised water (MilliQ, Millipore)
12. Vacuum centrifuge
13. 8M urea (Sigma Aldrich, U5128)
14. 10 mM Dithiolthreitol (DTT; Aldrich, 150460)
15. 20 mM Iodoacetamide (IAA; Sigma Aldrich, I6125)
16. 2 mL Zeba© desalt spin column (Thermo Scientific, 89890)
17. 50 mM Ammonium bicarbonate pH 8.6 (Fluka, purum p.a. ≥99.0%, 09832)
18. Peptide N-glycosidase F (PNGase F, Roche Diagnostics, 11 365 193 001)
19. Eppendorf tube incubator and shaker (Eppendorf, Thermomixer comfort)
20. Acetone (Sigma Aldrich, 270725)
21. 60% aqueous Methanol solution

Preparation of skim milk

1. Thaw human milk to 4 °C.

2. Add mammalian protease inhibitor cocktail to thawed human milk at 1 μL per mL.

3. Centrifuge milk at 10 000 rcf, 4 °C for 30 mins.

4. Freeze milk at 20 °C (to solidify cream).

5. Remove the solidified cream from skim milk by spatula

Precipitation and removal of casein

1. Thaw the skim milk to 4 °C and keep on ice.

2. Adjust the pH of the skim milk to 4.6 with 1M HCl.

3. Add $CaCl_2$ to the skim milk to a final concentration of 100 mM.

4. Ultracentrifuge the skim milk at 100 000 g, 4 °C for 90 minutes to pellet precipitated caseins.

5. Remove the casein-depleted skim milk fraction by pipette.

Isolation of skim milk proteins

1. Pipette 15 mL of casein-depleted skim milk into a 50 mL Falcon tube.

2. Add 10 mL of MeOH, vortex for 10 secs.

3. Add 10 mL of $CHCl_3$, vortex for 10 secs.

4. Add 15 mL of MilliQ water, vortex for 10 secs.

5. Centrifuge the mixture at 10 000 rcf, 4 °C for 15 mins (proteins will precipitate at the interphase).

6. Remove the upper aqueous phase (contains lactose, free oligosaccharides, soluble components).

7. To the precipitated proteins and the lower organic phase, add 15 mL of MeOH (reduces the density of the organic phase, allowing proteins to be collected by centrifugation).

8. Centrifuge the proteins and organic phase at 10 000 rcf, 4 °C for 10 mins to pellet the proteins.

9. Remove the organic phase by pipette.

10. Transfer the milk protein pellet to a 1.5 mL Eppendorf tube and by vacuum centrifugation (weigh the tube beforehand to obtain the dry weight of precipitated proteins).

In-solution release of *N*-linked glycans from skim milk proteins

1. Dissolve the milk protein pellet in 1 mL 8M urea (calculate mg/mL from the dry weight of the pellet, 8M urea can be avoided in the case of more soluble proteins).

2. Dilute a maximum of 10 mg of milk proteins in 250 µL of 8M urea (amounts > 10 mg can cause complications downstream).

3. Reduce the milk proteins with DTT (final conc. 10 mM) at 80 °C for 30 mins

4. Alkylate the reduced milk proteins with IAA (final conc. 20 mM) in the dark at room temperature for 30 mins.

5. Buffer exchange reduced and alkylated milk proteins using Zeba© desalt spin columns (must first be conditioned, buffer exchange can be avoided if 8M urea was not required to solubilise proteins).

6. Placing the Zeba© spin columns in a 15 mL Falcon tube. Condition the columns by loading 1 mL of 50 mM ammonium bicarbonate pH 8.6 and centrifuging at 1000 g for 2 mins to remove storage solution.

7. Wash the columns twice by loading 1 mL of 50 mM ammonium bicarbonate pH 8.6 and centrifuging at 1000 g for 2 mins.

8. Apply the reduced and alkylated milk proteins to the conditioned Zeba© desalt spin columns then centrifuge at 1000 g for 2 mins to buffer exchange proteins into 50 mM ammonium bicarbonate pH 8.6.

9. Transfer the proteins into a 1.5 mL Eppendorf tube.

10. Add PNGase F (10 units/mg), incubate at 37 °C for 24 hours with 500 rpm shaking.

Isolation of released *N*-linked glycans

1. To the 250 μL reaction volume add 1 mL of ice-cold acetone and leave at -20 °C for at least 1 hr (leave overnight for maximum precipitation of proteins and released *N*-linked glycans).

2. Centrifuge the acetone precipitation mixture at 3000 rcf for 3 mins (pellets both proteins and released *N*-linked glycans).

3. Removed the acetone supernatant making sure not to disturb the pellet.

4. To recover released *N*-linked glycans, re-suspended the pellet in 500 μL of 60% ice-cold aqueous then centrifuged at 3000 rcf for 3 mins to pellet the proteins (released *N*-linked glycans remain in the supernatant).

5. Repeat step 4 twice, combine the MeOH extract and dry by vacuum centrifuge.

Appendix II

Protocol for the Preparative Separation of N-linked Glycans by PGC-HPLC with Offline MALDI-TOF MS Detection

Materials

1. Deionised water (MilliQ, Millipore)

2. Ammonium acetate (NH_4COOCH_3; Fluka, 17836)

3. Sodium borohydride ($NaBH_4$; Aldrich, 213462)

4. Potassium hydroxide (KOH; Sigma, P5958)

5. Acetic acid (CH_3COOH; Sigma Aldrich, 320099)

6. Strata X polymeric reversed phase syringe columns (Phenomenex, 8B-S100-TBJ)

7. Dowex AG-50W-X8 cation exchange resin 300–1,180 µm wet bead size (Biorad, 142-1421)

8. HyperSep Hypercarb graphitised carbon solid phase extraction (SPE) column Acetonitrile (ACN; Sigma Aldrich, 34851)

9. Trifluoracetic acid (TFA; Sigma Aldrich, T6508)

10. HPLC system for binary gradient with a pressure limit of 150 bar or higher, containing an automatic sample injector, and an automatic fraction collector equipped with a 96-well plate holder (Shimadzu)

11. Hypercarb PGC column (2.1 mm × 100mm, 5µm, Thermo Scientific)

12. 2,5-dihydroxy benzoic acid (DHB; Sigma Aldrich, 85707)

13. Sodium acetate ($NaCOOCH_3$; Sigma Aldrich S-2889)

14. Vacuum centrifuge

15. MALDI-TOF mass spectrometer (Applied Biosystems 4800 *Plus* MALDI TOF/TOF Analyzer)

Reduction of Released Glycans

1. Dissolve the released N-linked glycans (Appendix I) in 100 μL of MilliQ water.

2. Add 20 μL of 100 mM NH$_4$COOCH$_3$ pH 5 and incubate at room temperature for 1 hr.

3. Add 120 μL of 2M NaBH$_4$ in 100 mM KOH and incubate at 50 °C for 3 hours.

4. Quench the reaction with the addition of 10 μL of glacial CH$_3$COOH.

5. Desalt the reduced N-linked glycans by solid phase extraction (SPE) with Dowex cation exchange resin and graphitised carbon.

Solid Phase Extraction with Dowex AG-50W-X8 Cation Exchange Resin

1. Prepare the Dowex cation exchange resin by suspending in 50% (v/v) MeOH

2. Wash with 50% (v/v) MeOH twice and store in 50% (v/v) MeOH

3. Place ~1 mL bed volume of Dowex cation exchange resin on a Strata X syringe column.

4. Condition the column by washing three times with 1 mL 1M HCl.

5. Wash three times with 1 mL MeOH.

6. Wash three times with 1 mL MilliQ water.

7. Transfer to a new tube, apply the reduced N-linked glycans (foaming is normal, do not load entire sample in one go, instead apply small aliquots to prevent excessive foaming).

8. Elute N-linked glycans with 1 mL MilliQ water.

9. Desalt further by SPE with HyperSep graphitised carbon.

Solid Phase Extraction with Graphitised Carbon

1. Condition the HyperSep graphitised carbon column by washing three times with 2 mL of 80% (v/v) ACN with 0.1% (v/v) TFA.

2. Wash three times with 2 mL of MilliQ water.

3. Apply the reduced *N*-linked glycans and wash three times with 2 mL of MilliQ water (to remove salts)

4. Transfer the column to a new tube and elute the neutral *N*-linked glycans with 20% (v/v) ACN.

5. Transfer the column to a new tube and elute the acidic *N*-linked glycans with 40% (v/v) ACN with 0.05% (v/v) TFA (alternatively both neutral and acidic *N*-linked glycans can be eluted together in this solvent).

6. Dry *N*-linked glycans by vacuum centrifuge.

Separation of *N*-linked Glycans by Porous Graphitised Carbon High Performance Liquid Chromatography (PGC-HPLC)

1. *Preparation of solvent A* (10 mM ammonium bicarbonate): Dissolve 7.906 g of ammonium bicarbonate into 1 L of MilliQ water to prepare a stock solution of 100 mM ammonium bicarbonate. Filter through a 0.45 μm filter. To 100 mL of stock solution (100 mM ammonium bicarbonate), add 900 mL of MilliQ water. Degas by in sonicator bath for 15 mins.

2. *Preparation of solvent B* (45% (v/v) ACN in 10 mM ammonium bicarbonate): Add 450 mL of ACN to 100 mL of stock solution (100 mM ammonium bicarbonate) and make up to 1 L with MilliQ water. Degas by in sonicator bath for 15 mins.

3. Set up HPLC system with automatic injector and fraction collector.

4. Wash system for 10 min at 100% solvent A at a flow rate of 1 mL/min (ensure the injection loop is washed with 100% solvent A).

5. Attach the Hyerpcarb PGC column 2.1 mm × 100mm, 5μm, and equilibrate for at least 30 mins at 100% solvent A at a flow rate of 250 μL/min.

6. Set up the following gradient: 100% solvent A for 15 mins, 0 − 5% solvent B over 0.1 mins, 5 − 35% solvent B over 44.9 mins, 35 − 100% solvent B over 20 mins, 100%

solvent B for 5 mins, 100 – 0% solvent B over 1 min then 0%B for 14 mins.

7. Collect fractions by time based collection at 1 min per fraction.

8. Dry fractions by vacuum centrifugation.

Matrix Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry (MALDI-TOF-MS)

1. *Preparation of MALDI matrix*: Dissolve 10 µg/µL of DHB in 70% (v/v) with 10 mM sodium acetate.

2. *Sample preparation*: Apply 1 µL of DHB solution to a spot on the MALDI target. Apply 1 µL of sample to droplet and allow to air dry.

3. *Instrument parameters*: Analyse neutral *N*-linked glycans in reflectron-positive ion mode.

Appendix III

MALDI-TOF Mass Spectra of Fractions Separated by PGC-HPLC

Fraction 49

4700 Reflector Spec #1[BP = 603.9, 3386]
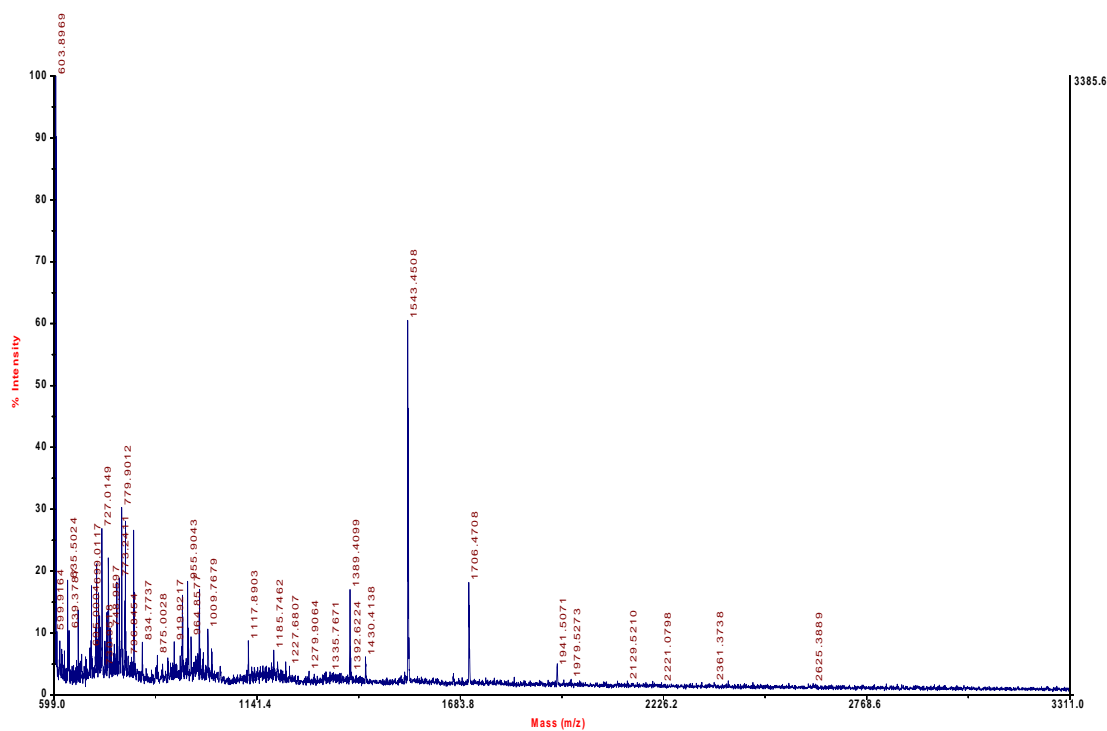
Fraction 50

4700 Reflector Spec #1[BP = 1706.4, 881]
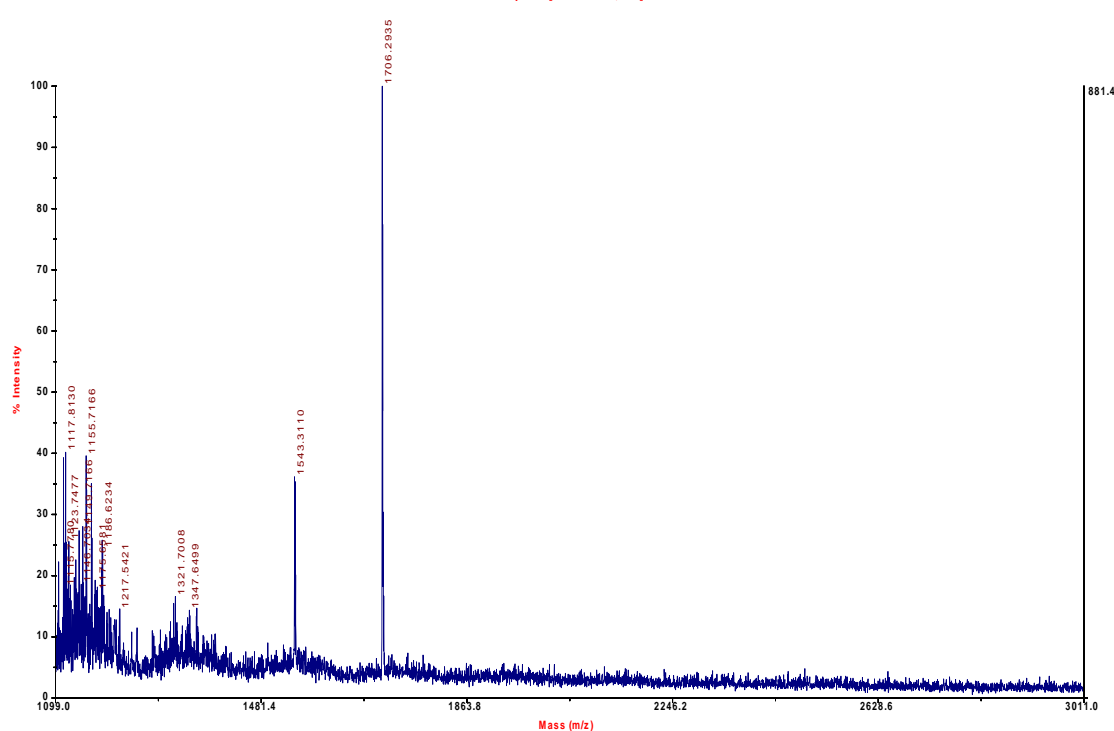
Fraction 51
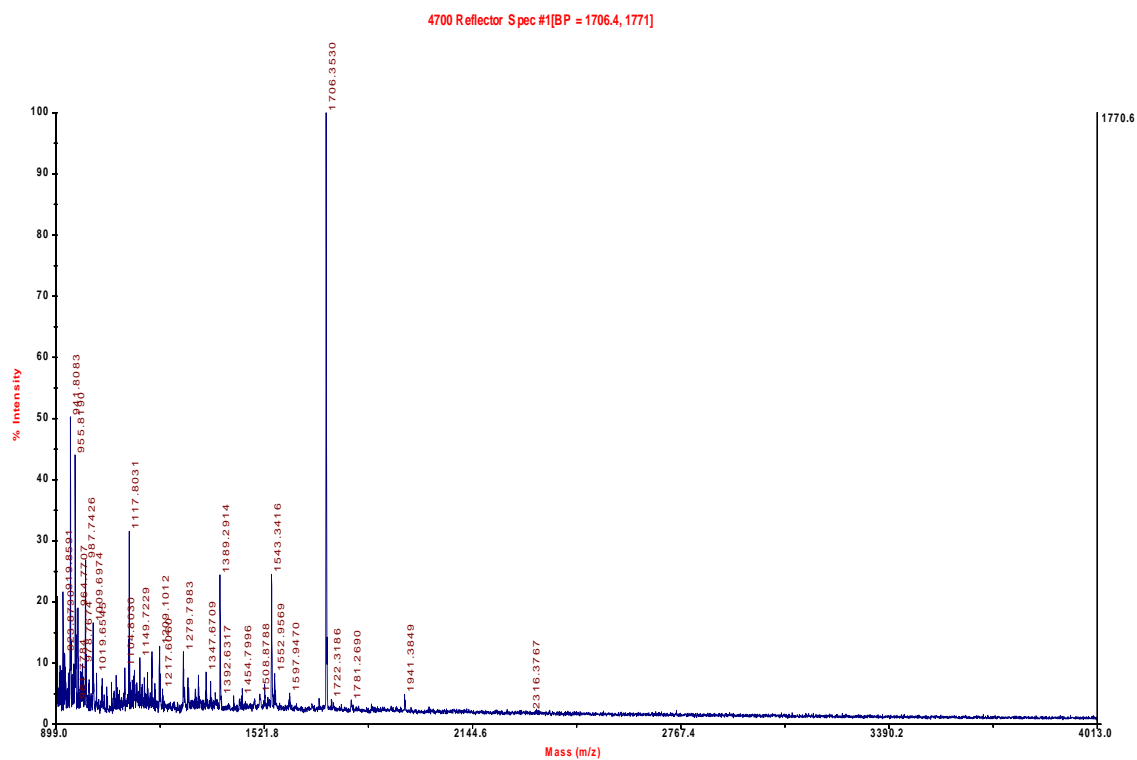
Fraction 52
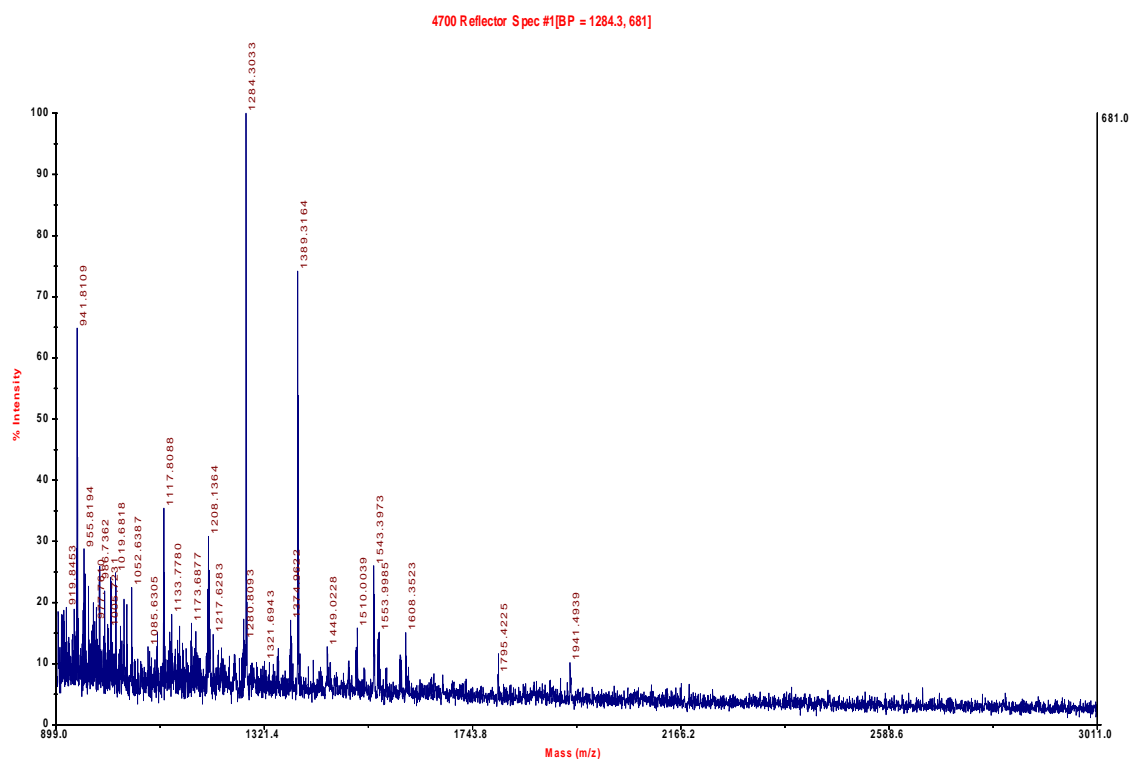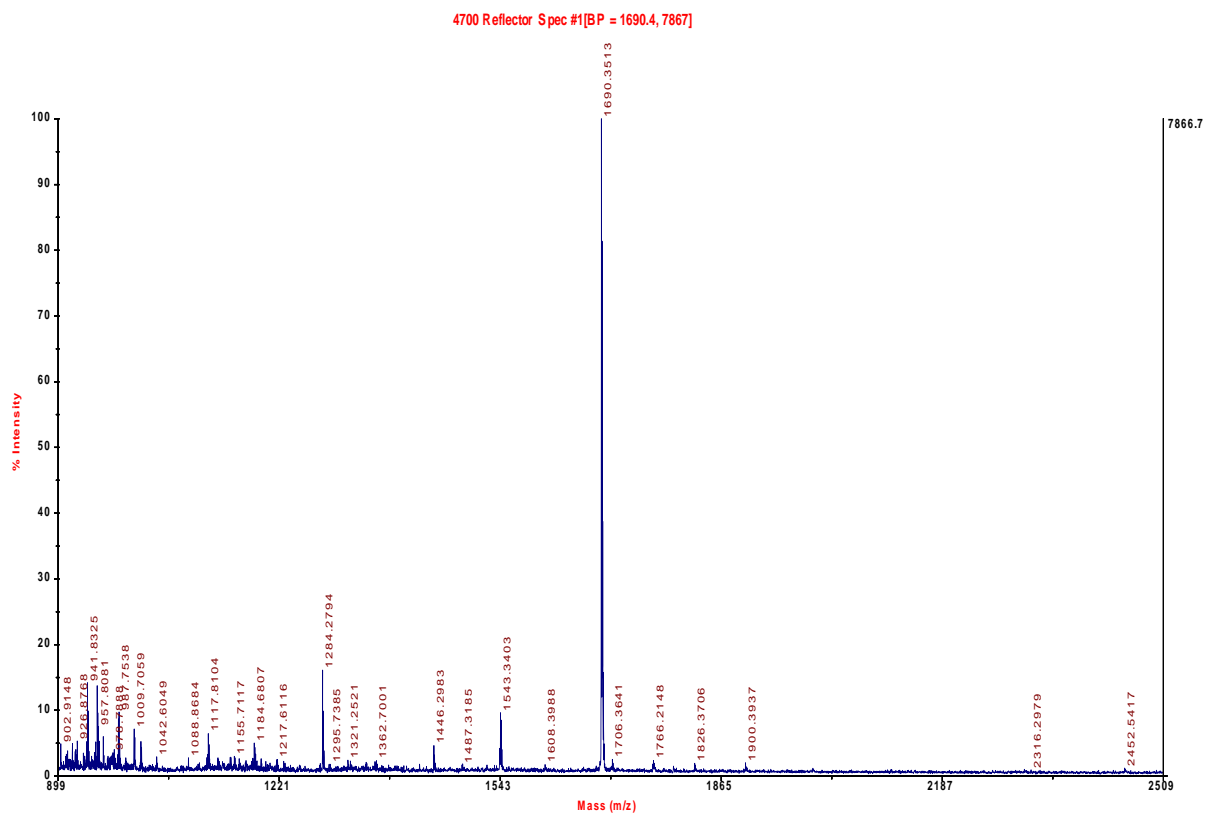
152

Fraction 53

Fraction 54
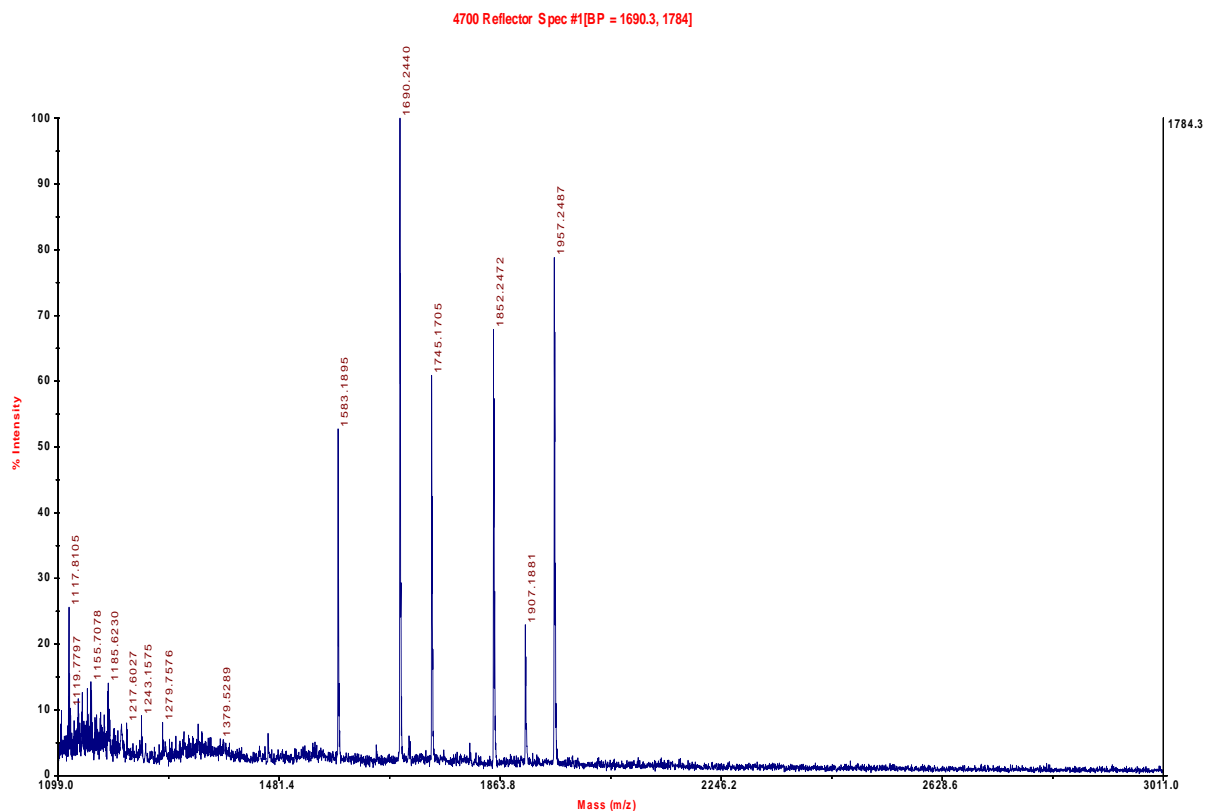
153

Fraction 55

Fraction 56

154
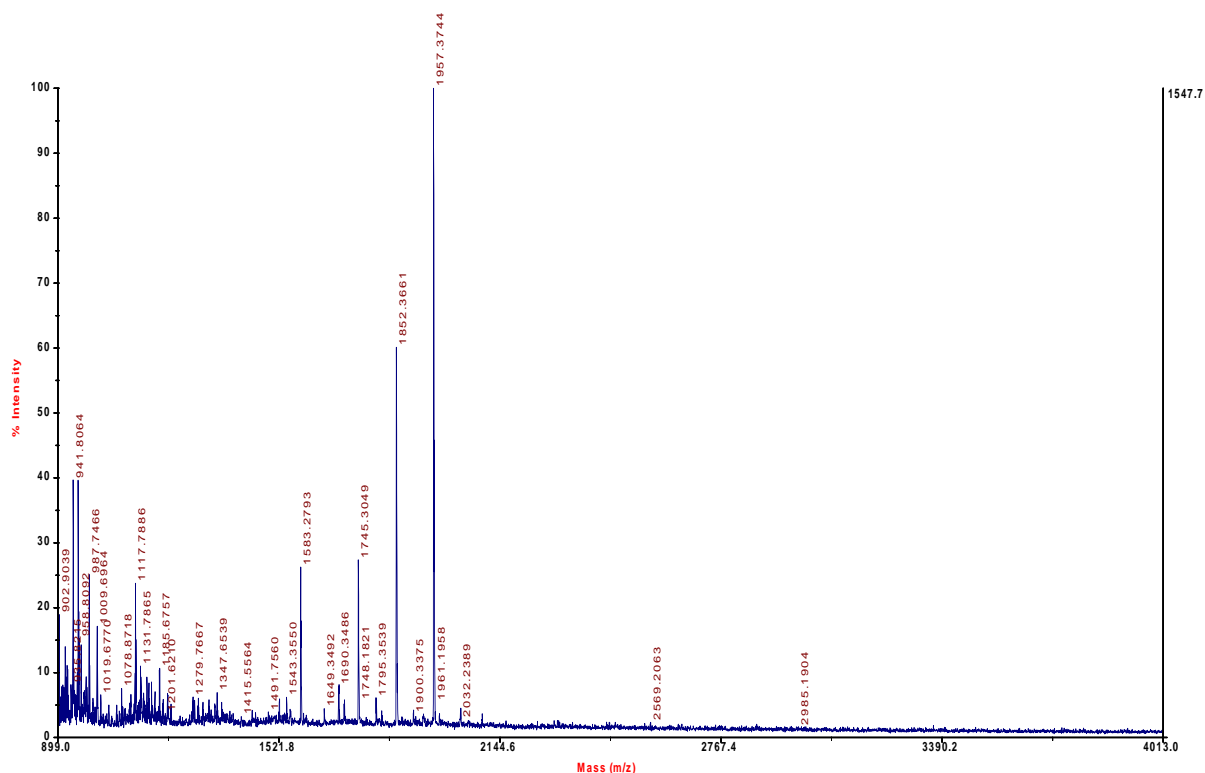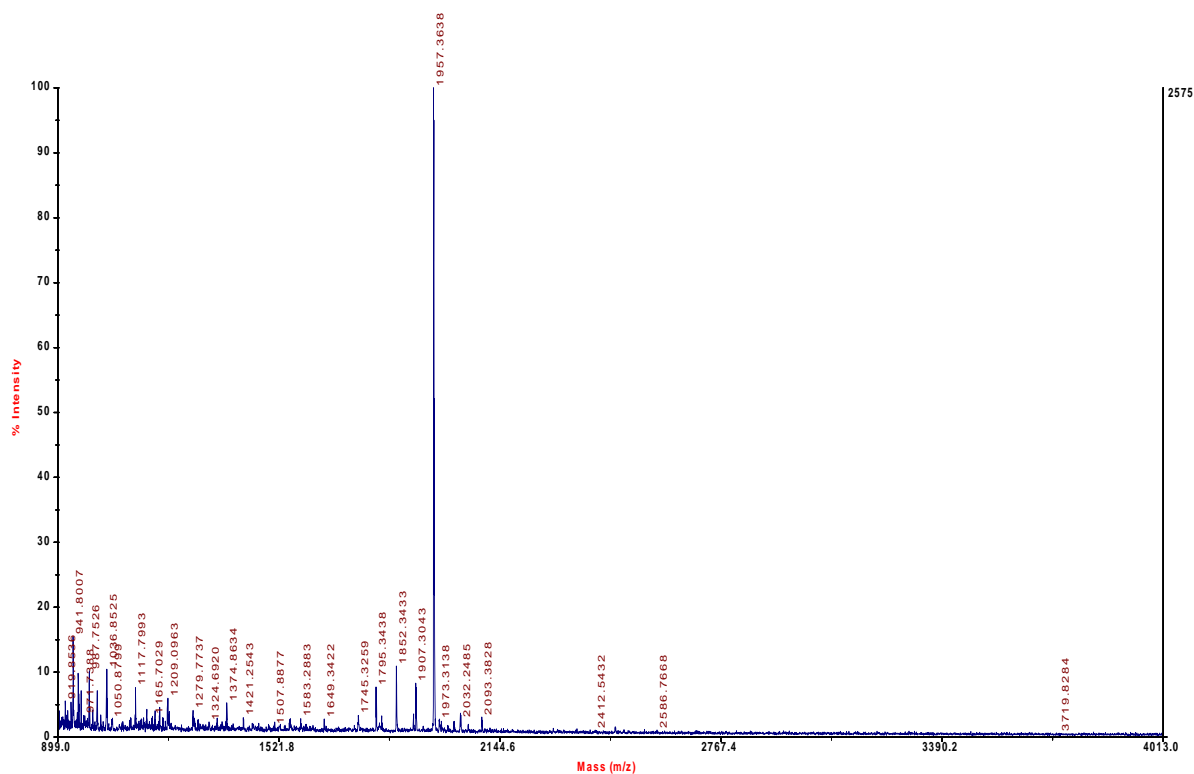
Fraction 57

Fraction 58

Fraction 59

Fraction 60

Fraction 61

Fraction 62

157

Fraction 63

Fraction 64

158

Fraction 65

Fraction 66

159

Fraction 67