Imperial College London

Mass spectrometric investigation of biomedically important glycosylation

A thesis submitted for the Degree of Doctor of Philosophy of Imperial College London

Submitted by

Qiushi Chen

Imperial College London

Department of Life Sciences

South Kensington

London SW7 2AZ

United Kingdom

Abstract

Glycobiology is the comprehensive study of the structure, biosynthesis, function and evolution of saccharides which are also named sugars or glycans. Glycosylation is a type of modification in which sugars are added to another molecule, such as a protein molecule or a ceramide. Abnormal glycosylation is frequently associated with diseases such as cancer and immune responses. Defining glycan structures is therefore important for understanding glycan function in health and disease. In addition, identification of glycan populations can provide essential information for further research on glycoproteins and glycolipids. In this thesis, glycomic experimental approaches were employed to characterize the structures and populations of glycans of glycoconjugates from HeLa cells, normal human dermal fibroblast (NHDF) cells, myoblasts, myotubes and trophoblasts. These approaches include sample preparation methodologies which were followed by the application of highly sensitive mass spectrometry, particularly MALDI-TOF MS, MALDI-TOF/TOF MS/MS and GC-MS.

Ribosome inactivating proteins (RIPs) and lectins from elderberry are more toxic to HeLa cells than to NHDF cells. The difference in the cytotoxicity was hypothesized to be caused by the difference in the glycome patterns of HeLa and NHDF cells. To test the hypothesis, glycome patterns on both glycoproteins and glycolipids of HeLa and NHDF cells were investigated. Glycomic results have revealed that glycome patterns in HeLa cells and NHDF are different, and this gives a possible explanation for the difference observed in the cytotoxicity assay.

Glutamine-fructose-6-phosphate transaminase 1 (GFPT1) is the first enzyme of the hexosamine biosynthetic pathway which yields uridine diphosphate N-acetylglucosamine (UDP-GlcNAc), an essential substrate for protein glycosylation. N-glycan branching is especially sensitive to alterations in the concentration of this sugar nucleotide. Mutations in the gene *GFPT1* can result in "limb-girdle CMS with tubular aggregates" which is a subtype of congenital myasthenic syndromes (CMS). To investigate whether protein glycosylation at the neuromuscular junction might be involved in this impairment, the N-glycomes of myoblasts and myotubes derived from healthy controls and patients were investigated. My result showed that global glycosylation is not significantly impaired in the muscle cells from the CMS patients caused by *GFPT1* mutations.

The human fetoembryonic defense system hypothesis (hu-FEDS) is a hypothetical model depicting a way via which the human immune system is able to recognize foreign substances as "own species" as has been observed with maternal immune tolerance in pregnancy. The fundamental idea of this hypothesis is that glycoproteins existing in the reproductive system and exposed on gametes can either inhibit immune responses or prevent rejection of the foetus. This model has not been tested in human trophoblasts. My glycomic analyses of three trophoblast populations (CTB, STB and evCTB) revealed that functional glycan structures that are present on human gametes are also expressed on trophoblasts, and this provides further evidence for the hu-FEDS hypothesis.

Content

List of Figures

List of Tables

List of Abbreviations

AChE	Acetylcholinesterase
ALG	Asparagine-linked glycosylation
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
Cer	Ceramide
CMS	Congenital myasthenic syndrome
CRD	Carbohydrate recognition domain
CTB	Cytotrophoblast
CTL	Cytotoxic T lymphocyte
Cys	Cysteine
Da	Dalton, mass unit
DABP	Diaminobenzophenone
DHB	Dihydroxybenzoic acid
DTT	Dithiothreitol
eu-FEDS	eutherian fetoembryonic defence system hypothesis
evCTB	extravillous cytotrophoblast
EI	Electron ionization
E-PHA	Erythroagglutinating phytohemagglutinin
ER	Endoplasmic reticulum
FAB	Fast atom bombardment
FCS	Foetal calf serum
Fuc	Fucose
Fuc-T	Fucosyltransferase
Gal	Galactose
GalNAc	N-acetylgalactosamine
GalNAc-T	N-acetylgalactosaminyltransferase
Gal-T	Galactosyltransferase

GdA	Glycodelin-A
GDP	Guanosine diphosphate
GFPT1	Glutamine-Fructose-6-Phosphate Transaminase 1
Glc	Glucose
GlcCer	Glucosylceramide
GlcNAc	N-acetylglucosamine
GlcNAc-T	N-acetylglucosamine transferase
Glc-T	Glucosyltransferase
Gln	Glutamine
GSL	Glycosphingolipid
hu-FEDS	human fetoembryonic defense system hypothesis
Hex	Hexose
HexNAc	N-acetylhexosamine
His	Histidine
HLA	Human leukocyte antigens
IAA	Iodoacetic acid
Ile	Isoleucine
keV	kilo electron volt
kV	kilovolts
KBH_4	Potassium borohydride
KOH	Potassium hydroxide
LacNAc	N-acetyllactosamine
LGMD2A	Limb girdle muscular dystrophy type 2A
m/z	mass to charge ratio
MALDI	Matrix Assisted Laser Desorption Ionization
Man	Mannose
Met	Methionine
MS	Mass Spectrometry
MS/MS	Tandem mass spectrometry
MTND5	Mitochondrial NADH dehydrogenase 5
$NaBD_4$	Sodium borodeuteride

NeuAc	N-acetylneuraminic acid
NHDF	Normal human dermal fibroblast
NK cell	Natural killer cell
NMJ	Neuromuscular junction
OST	Oligosaccharyltransferase
pp-GalNac-T	Polypeptide N-acetylgalactosaminyltransferase
P	A phosphate group
Phe	Phenylalanine
PNGase F	Peptide N-Glycosidase F
Pro	Proline
PDB	Protein Data Bank
PP13	Placental protein 13
RI	Relative intensity
RIP	Ribosome inactivating protein
Sia-T	Sialyltransferase
Ser	Serine
SIS	Sugar isomerase
SNA	Sambucus nigra agglutinin
STB	Syncytiotrophoblast
S/N	Signal to noise ratio
TFA	Trifluoroacetic acid
Thr	Threonine
TOF	Time-of-flight
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
Tyr	Tyrosine
U	Unit
UDP	Uridine diphosphate
Val	Valine
^{o}C	Degree centigrade

Declaration

I hereby make a declaration that the work presented in this thesis has not been previously or concurrently submitted for any other degree, diploma or other qualification at other university or college, and it is the result of my own independent investigation unless otherwise stated.

Qiushi Chen

Acknowledgements

The research study presented in the thesis has been carried out in the laboratory of Professor Anne Dell and Dr. Stuart M. Haslam in the Department of Life Sciences at Imperial College London. I am grateful to my supervisors, Professor Anne Dell and Dr. Stuart M. Haslam who gave me a chance to work in the laboratory. I am proud of being supervised by the two most prestigious supervisors that anyone could ever have.

Although a thousand words cannot convey a sufficient gratitude to my supervisors, I still would like to express my gratitude to them, whose guidance and knowledge has contributed greatly to my project. I feel very lucky to be their student. Without the help from them, my projects would not be possible and most importantly I would not be able to sit here and write this acknowledgement. Their rigorous scientific training and meticulous scrutiny have influenced me deeply and will continue helping me for the rest of my life.

'桃花潭水深千尺，不及汪伦送我情'(Chinese phonetic alphabet version: tao hua tan shui shen qian chi, bu ji wang lun song wo qing) is a verse written by Li Bai who is a great Chinese poet considered from the flourishing of Chinese poetry in the mid-Tang Dynasty to the present as a genius who took traditional poetry to a new height. This verse is a metaphor indicating that the help and support that the author's friend provided for the author is even deeper than the depth of Peach Blossom Pool which is one thousand feet. Now my gratitude can also be expressed by this verse.

I should also acknowledge to Dr. Paola Grassi, Dr. Poh-Choo Pang, Dr. Aristotelis Antonopoulos and Gang Wu (now he is Dr. Wu) who have assisted me in doing experiment and analysing data. Thanks to Dr. Simon North who has spent his precious time on proofreading this thesis. Many thanks also go to other laboratory members who have helped me directly or indirectly throughout this 3-4 year period: Professor Howard Morris, Dr. Maria Panico, Dr. Paul Hitchen, Dr. Kevin Canis, Dr. David Damerell, Lesley Harris, Federico Sastre, Dinah Rahman, Valeria V. Ventura (now she is Dr. Ventura), Nan Jia, Grigorij Sutov, Tiandi Yang and Laura Bouche. Without their support and friendship, the difficulty would be more difficult and the depressing moment would be much tougher.

In addition, thanks to my collaborators Professor Hanns Lochmuller and Dr. Juliane Mueller (Newcastle University, UK) for providing *GFPT1* myoblast and myotube samples, thanks to

Thesis publications

SHANG, C., **CHEN, Q.**, DELL, A., HASLAM, S. M., DE VOS, W. H. & VAN DAMME, E. J. 2015. The Cytotoxicity of Elderberry Ribosome-Inactivating Proteins Is Not Solely Determined by Their Protein Translation Inhibition Activity. *PLoS One*, 10, e0132389.

CHEN, Q., MULLER, J.S., PANG, P. C., LAVAL, S.H., HASLAM, S.M., LOCHMULLER, H., DELL, A. 2015. Global N-linked Glycosylation is Not Significantly Impaired in Myoblasts in Congenital Myasthenic Syndromes Caused by Defective Glutamine-Fructose-6-Phosphate Transaminase 1 (GFPT1). *Biomolecules*, 5, 2758-2781.

CHEN, Q., PANG, P. C., COHEN, M. E., LONGTINE, M. S., SCHUST, D. J., HASLAM, S. M., BLOIS, S. M., DELL, A., CLARK, G. F. 2015. Evidence for Functional Glycosylation of Human Trophoblasts. *Molecular & Cellular Proteomics*. Submitted.

Chapter 1

Introduction

1. Introduction

1.1. Glycobiology

Glycobiology is the comprehensive study of the structure, biosynthesis, function and evolution of saccharides in biological systems (Kiessling and Splain, 2010; Cummings and Pierce, 2014). It has only recently become a branch of molecular and cellular biology. The term glycobiology, formally coined in August 1988, is used to recognize the coming together of the traditional disciplines of carbohydrate chemistry and biochemistry (Blow, 2009). This union is a consequence of the dramatically increased understanding of glycans in molecular and cellular biology in the last 30 years.

Glycans are compounds consisting of a series of glycosidically linked monosaccharides. Like amino acids and nucleic acids, glycans are also widely distributed in nature (Hart, 2013). All cells in nature are covered with a complex array of sugar chains, termed the glycocalyx (Varki, 2007). As an example, an erythrocyte surface glycocalyx in cross-section is shown in Figure 1.1. Consistent with the high glycan abundance on cells, genomic sequencing studies from eubacteria to eukaryotes show that approximately 1% of the genome contributes to sugar processing enzymes. Moreover, these glycan related genes are highly conserved (Coutinho et al., 2003; Kiessling and Splain, 2010).
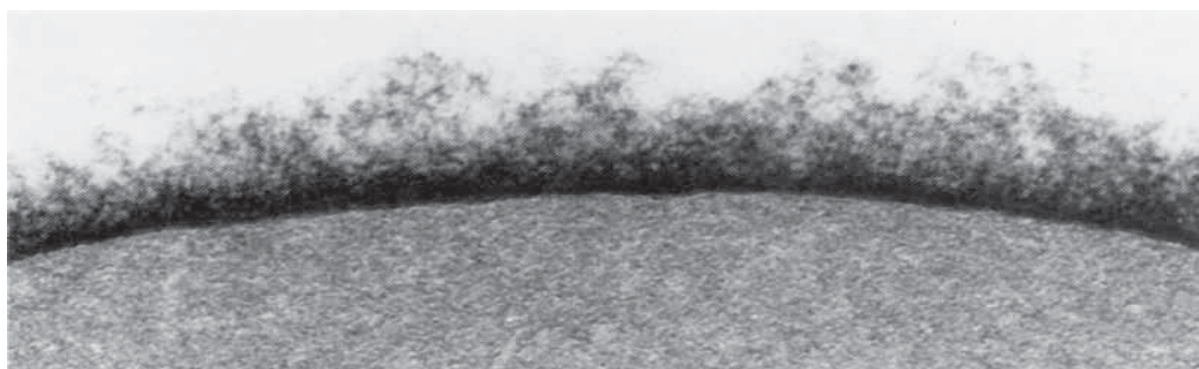
Figure 1.1 Electron microscopy: an erythrocyte surface glycocalyx

The cell has been stained using special staining techniques. It is up to 1400 Å thick, and the oligosaccharide filaments are 12–25 Å in diameter. The glycocalyx is comprised of abundant carbohydrate-rich molecules, polysaccharides, proteoglycans, glycoproteins and glycolipids (Roseman, 2001; Voet and Voet, 2004).

It has long been known that carbohydrates serve as an energy source. However, a re-evaluation of the biological role of glycans is required (Rademacher et al., 1988) because glycans have many other functions. For example, they were found to be the essence of antigenic determinants in the ABO blood group system (Cohen et al., 2009). In addition, glycans play a pivotal role in the attachment of the influenza virus to the host cells (Chu and Whittaker, 2004; Varki, 2007). Indeed glycans are involved in the recognition of host cells by most pathogens. Over the last 30 years a large amount of data has been assembled demonstrating the importance of glycans. In point of fact, for the most part, carbohydrates do not exist and function as simple sugars but as complex molecular conjugates. Thus glycans are usually found in covalent association with proteins and lipids via glycosylation (Rademacher et al., 1988).

1.2. Glycosylation

In biology, glycosylation is an enzymatic reaction that attaches glycans to proteins, lipids, or other organic molecules (Taylor and Drickamer, 2011; Freeze, 2006). In this thesis, the main focus is protein N- and O-glycosylation and glycosphingolipid glycosylation.

1.2.1. Monosaccharides

Monosaccharides are the basic constituents of glycans. Hexoses are monosaccharides which possess six carbons designated carbon 1 (C1) to carbon 6 (C6) from the top which is an aldehyde group to the bottom which is a hydroxymethyl group (Figure 1.2). Four of these six carbons (C2, C3, C4 and C5) are chiral, which means each one has two configurations, so theoretically there are sixteen (2^4) potential structures, half of which are defined as L-hexoses and the remaining are termed D-hexoses. The difference between the D- and L- structure is determined by the configuration of C5. The structure of a D-monosaccharide is the mirror image of the structure of its corresponding L-form. Two stereoisomers (not identical) that are mirror images of each other are termed enantiomers. For instance, D-Glucose and L-Glucose are enantiomers. An alteration of configuration at each chiral carbon is called epimerization,

which will yield a new monosaccharide. For instance, epimerization at C4 of D-Glucose produces D-Galactose (Figure 1.2), and thus these two monosaccharides are termed epimers.
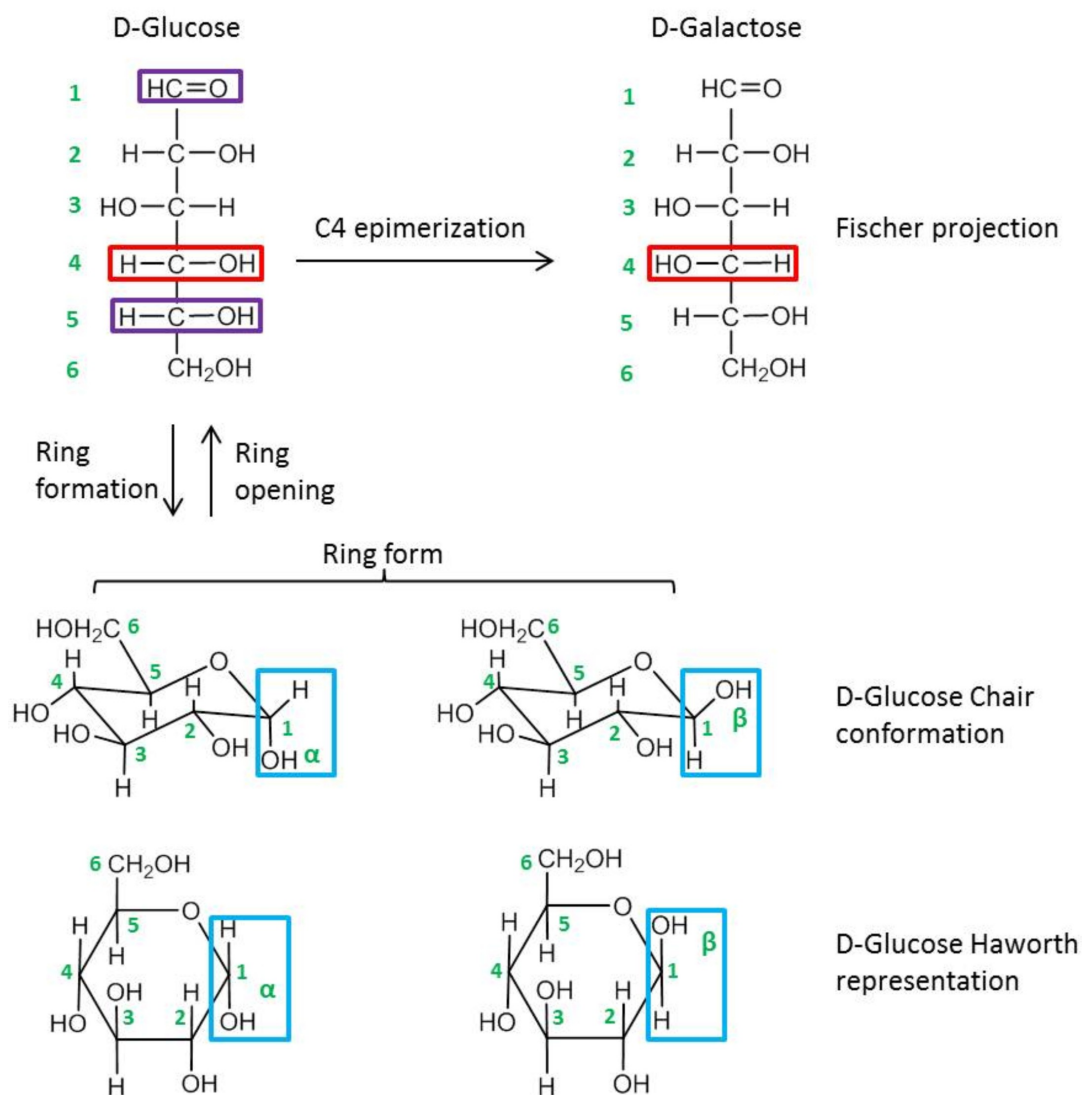
Figure 1.2 Fischer projection of D-Glucose and D-Galactose, chair conformations of D-Glucose and Haworth representations of D-Glucose

The D-Glucose linear structure is shown in the form of Fischer projection. It converted into D-Galactose after the epimerization occurred at C4 (red frame) of the Glu. The ring structure of D-Glucose is formed via the reaction of the hydroxyl group at C5 and the aldehyde group at C1 which are labelled using purple frames. C1 in α and β anomers of D-Glucose are emphasized in light blue frames.

The hexose linear structure is transformed into the ring structure via reaction of the hydroxyl group at C5 with aldehyde group at C1 (Figure 1.2). In the ring structure, C1 becomes a chiral or "anomeric" carbon, and the different configurations of this carbon are known as α and β anomers. It is easy to distinguish these two anomers: in the Haworth representation of a

D-monosaccharide if the anomeric hydroxyl group is below the ring plane the monosaccharide is the α anomer; if the hydroxyl group is above the plane the monosaccharide is the β anomer.

Other monosaccharides can be obtained from hexoses via some modifications. For instance, a deoxyhexose can be acquired via the removal of an oxygen from a hexose, and an N-acetylhexosamine (HexNAc) can be produced by replacing a hydroxyl group with an acetylated amino group. In addition, HexNAcs can be further modified to yield acidic sugars.

Monosaccharides are connected via glycosidic bonds to form oligo- and polysaccharides. Figure 1.3 displays a linkage between a glucose (Glc) and a galactose (Gal), which is formed via the reaction of the anomeric carbon C1 of the Gal and the hydroxyl group at C4 of the Glc. This disaccharide is called lactose. Because the oxygen in the hydroxyl group at the C1 position of Gal is in the β configuration and this oxygen is linked to the C4 of Glc, the glycosidic bond is called a β1,4 linkage. In this newly formed disaccharide, there is a reducing end and a non-reducing end. The reducing end is actually a hemiacetal group, which can reduce inorganic ions like Cu^{2+} while the non-reducing end does not have this activity. It is conventional that the non-reducing end and the reducing end are drawn on the left hand side and right hand side, respectively.
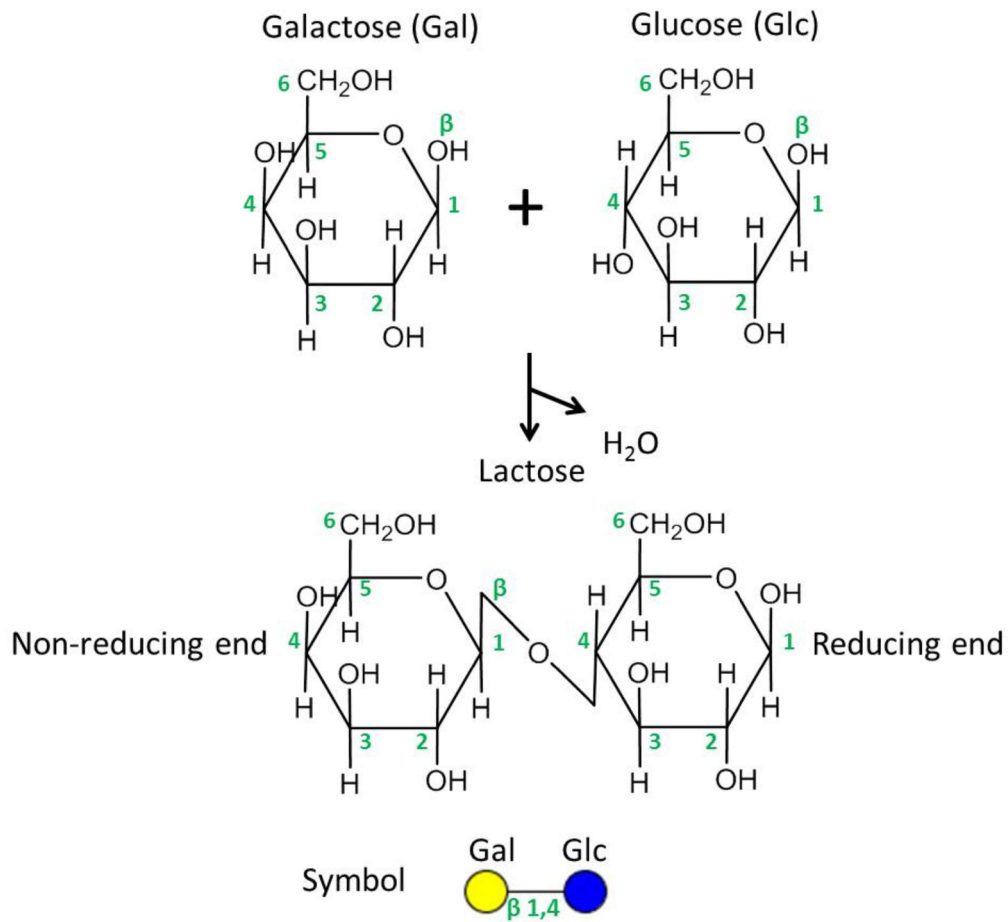
Figure 1.3 The β1,4 linkage between a Gal and a Glc

A β1,4 linkage is formed between a Gal and a Glc via the reaction of the anomeric carbon C1 of Gal and the hydroxyl group on the C4 of Glc. The product is lactose. The reducing end and the non-reducing end of the project are labelled. Cartoon symbols for Gal and Glc are also shown.

In addition, it is convenient to show a glycan, such as lactose, via symbolic representation; symbols such as these are used to represent the sequences of glycans in this thesis and are discussed in detail below.

Table 1.1 Seven common monosaccharides in humans

The classification, name, abbreviation and structure of each monosaccharide are listed.

Neutral monosaccharides			
D-Glucose (Glc)	D-Mannose (Man)	D-Galactose (Gal)	L-Fucose (Fuc)
Hexose			Deoxyhexose
(structure of D-Glucose) CH_2OH, O, OH, H, OH, HO, H, H, OH	(structure of D-Mannose) CH_2OH, O, OH, H, OH, OH, HO, H, H, H	(structure of D-Galactose) CH_2OH, O, OH, OH, H, OH, H, H, H, OH	(structure of L-Fucose) H, O, OH, H, CH_3, H, OH, HO, OH, H

Amino monosaccharides		Acidic monosaccharide
D-N-acetylglucosamine (GlcNAc)	D-N-acetylgalactosamine (GalNAc)	N-acetylneuraminic acid (NeuAc)
N-acetylhexosamine (HexNAc)		Sialic acid
(structure) CH_2OH, O, OH, H, OH, H, HO, H, H, $NHCOCH_3$	(structure) CH_2OH, O, OH, OH, H, OH, H, H, H, $NHCOCH_3$	(structure) H, O, COOH, R_2, R_1, H, H, H, OH, OH, H — $R_1 = CHOHCHOHCH_2OH$, $R_2 = CH_3CONH$

In human glycoproteins and glycolipids there are 7 common monosaccharides (Table 1.1): 4 neutral monosaccharides, 2 amino monosaccharides and 1 acidic monosaccharide (Goldberg et al., 2009). As shown in the table, Man and Gal are different from Glc at C2 and C4 respectively. Fucose (Fuc) is different from other monosaccharides as its natural configuration is the L-stereoisomer while others are D-form. The difference between D-Gal and D-Fuc is that at C6 position there is a hydroxyl group in the former while in the latter the hydroxyl group is replaced by a hydrogen atom, and thus the Fuc can also be called 6-deoxy-Gal. Replacing the hydroxyl group with an acetylated amino group at C2 in Glc and Gal forms N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc) respectively; both of which are amino monosaccharides. The acidic sugar (sialic acid) is more complicated: 1. it possesses 11 carbon atoms, 2. it still forms a ring structure, but unlike previously mentioned monosaccharides, the ring structure is formed via the reaction of the carbonyl

group at C2 and hydroxyl group at C6, 3. it carries a carboxyl group at C1, an acetylamino group at C5 and a glycerol group at C6.

Human glycoproteins have a limited repertoire of sugars; useful parameters of these glycans have been summarized in Table 1.2, which can be used as a reference for the glycomic data analysis.

Table 1.2 Parameters of the seven common monosaccharides in humans

The classification, abbreviation, symbol, mass, residue mass and permethylated residue mass of each monosaccharide are listed. The symbol is assigned to each monosaccharide according to the rules adopted by the Consortium for Functional Glycomics (CFG) (http://www.functionalglycomics.org/) and the Essentials of Glycobiology (http://www.ncbi.nlm.nih.gov/books/NBK1931/figure/ch1.f5/?report=objectonl) online textbook.

Monosaccharide	Symbol	Mass	Residue mass	Permethylated residue mass
Deoxyhexose	Fuc ▲	164	146	174
Hexose	Glc ● Gal ○ Man ●	180	162	204
HexNAc	GalNAc □ GlcNAc ■	221	203	245
Sialic acid	NeuAc ◆	309	291	361

1.2.2. Protein glycosylation

Protein glycosylation is ubiquitous in eukaryotes especially in extracellular matrices and cellular surfaces (Weerapana and Imperiali, 2006). It has been estimated that more than 50% of all proteins in nature are glycosylated (Apweiler et al., 1999). Glycosylation usually contributes to the maintenance of a protein structure; and therefore has an influence on the glycoprotein function. For instance, glycans contribute to the structure and stability of immunoglobulins, which is essential for their binding to receptors (Arnold et al., 2007; Schroeder and Cavacini, 2010). Sometimes, glycosylation-associated function is more dependent on the presence of the carbohydrates, rather than on the protein structure (Kent,

2004). For instance, the binding of human sperm to an ovum is regulated by a tetrasaccharide sialyl-Lewis X [NeuAcα2,3Galβ1,4(Fucα1,3) GlcNAc] on the zona pellucida (Pang et al., 2011).

As previously mentioned, glycans are attached to proteins by covalent bonds via glycosylation. There are two common glycosylation types: 'N-linked' to asparagine (Asn) residues and 'O-linked' to serine (Ser) or threonine (Thr) residues (Tissot et al., 2009; Taylor and Drickamer, 2011).

1.2.2.1.   N-glycosylation

Most N-glycosylation occurs co-translationally, which means the glycosylation occurs before the protein is folded (Culyba, 2012). Intriguingly, unlike proteins which use DNA as template, no template is involved for the biosynthesis of N-glycans (Kiessling and Splain, 2010).
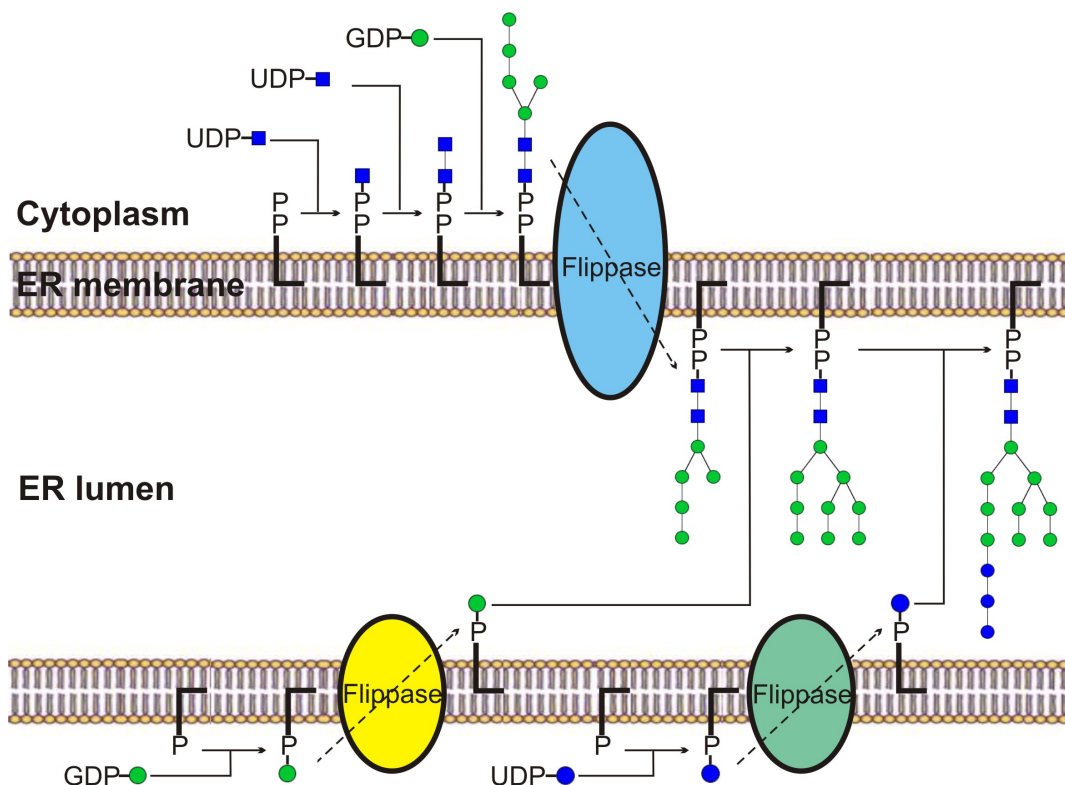
Figure 1.4 Biosynthesis of the lipid-linked N-glycan precursor
L, dolichol; PP, a pyrophosphate; ■ GlcNAc, ● Man, ● Glc.

The biosynthesis pathway of N-glycans can be divided into 3 steps:

1. A lipid-linked 14-sugar (2 GlcNAc, 9 Man and 3 Glc) precursor is produced by a series of glycosyltransferases (Kornfeld and Kornfeld, 1985; Freeze, 2006). The lipid and the oligosaccharide are connected via a pyrophosphate linkage. The synthesis of the precursor occurs in two cellular locations (Figure 1.4): firstly on the cytoplasmic side of the endoplasmic reticulum (ER) membrane, the first GlcNAc from a uridine diphosphate (UDP)– GlcNAc is attached to a dolichol via an enzyme, dolichyl-phosphate (UDP-N-acetylglucosamine) N-acetylglucosaminephosphotransferase 1 (DPAGT1), then the remaining 6 monosaccharides (1 GlcNAc and 5 Man) are synthesized using UDP–GlcNAc and guanosine diphosphate (GDP)–Man, after that this structure is translocated into the ER lumen via a flippase; secondly in the ER lumen another 7 monosaccharides (4 Man and 3 Glc) from dolichol phosphate Man and dolichol phosphate Glc are sequentially added to complete the synthesis of the precursor (Freeze et al., 2014). Defects in enzymes which are responsible for the biosynthesis of the 14-sugar (2 GlcNAc, 9 Man and 3 Glc) lipid-linked precursor oligosaccharide typically result in type 1 congenital disorders of glycosylation (CDG-1) (Freeze and Aebi, 2005; Freeze, 2006; Freeze et al., 2014).

2. The oligosaccharide is entirely transferred to an N-glycosylation site of a polypeptide by the oligosaccharyltransferase (OST) in the ER (Figure 1.5A). In mammalian cells, it starts with the formation of a β-linkage between the oligosaccharide reducing terminal N-acetylglucosamine (GlcNAc) and the amide nitrogen of asparagine (Asn) on a target polypeptide (Figure 1.5B) (Freeze, 2006; Lowe and Marth, 2003). The Asn is usually found in a conserved sequence Asn-B-Thr or Asn-B-Ser, in which B can be any amino acid except Proline (Pro), although Asn in other sequences, such as Asn-B-Cys (cysteine), can be glycosylated as well (Matsui et al., 2011). Pro cannot be the middle residue because it will prevent the formation of a loop which is essential for the glycan transfer.

3. Following the transfer of the precursor to the target polypeptide, oligosaccharide processing is carried out initially in the ER and then in the Golgi apparatus (Aebi et al., 2010). The three Glc residues are firstly removed; the outer one is cleaved by ER glucosidase I and the inner two are removed by ER glucosidase II. The remaining oligosaccharide is then processed by a series of mannosidases (ER mannosidase, Golgi mannosidases 1A, 1B and 1C) yielding a high mannose oligosaccharide (usually possessing 9 to 5 Man residues). The initiation of a hybrid or complex N-glycan begins with the action of GlcNAc transferase I

(GlcNAc-T I) on a heptasaccharide in the Golgi apparatus, the structure of which is Manα1,3(Manα1,3(Manα1,6Manα1,6))Manβ1,4GlcNAcβ1,4GlcNAc. After this, an octasaccharide is yielded. Hybrid glycans will be formed when mannosidases do not act on this octasaccharide. The two outer mannose residues in the octasaccharide are removed by α-Mannosidase II generating a substrate for GlcNAc-T II (Figure 1.5A). A Fuc can be transferred from GDP-Fuc to the proximal GlcNAc via the action of fucosyltransferase VIII. The resulting N-glycan is extended by the addition of Gal to generate a complex N-glycan (Freeze et al., 2014; Schwarz and Aebi, 2011; Freeze, 2006; Takahashi et al., 2009).
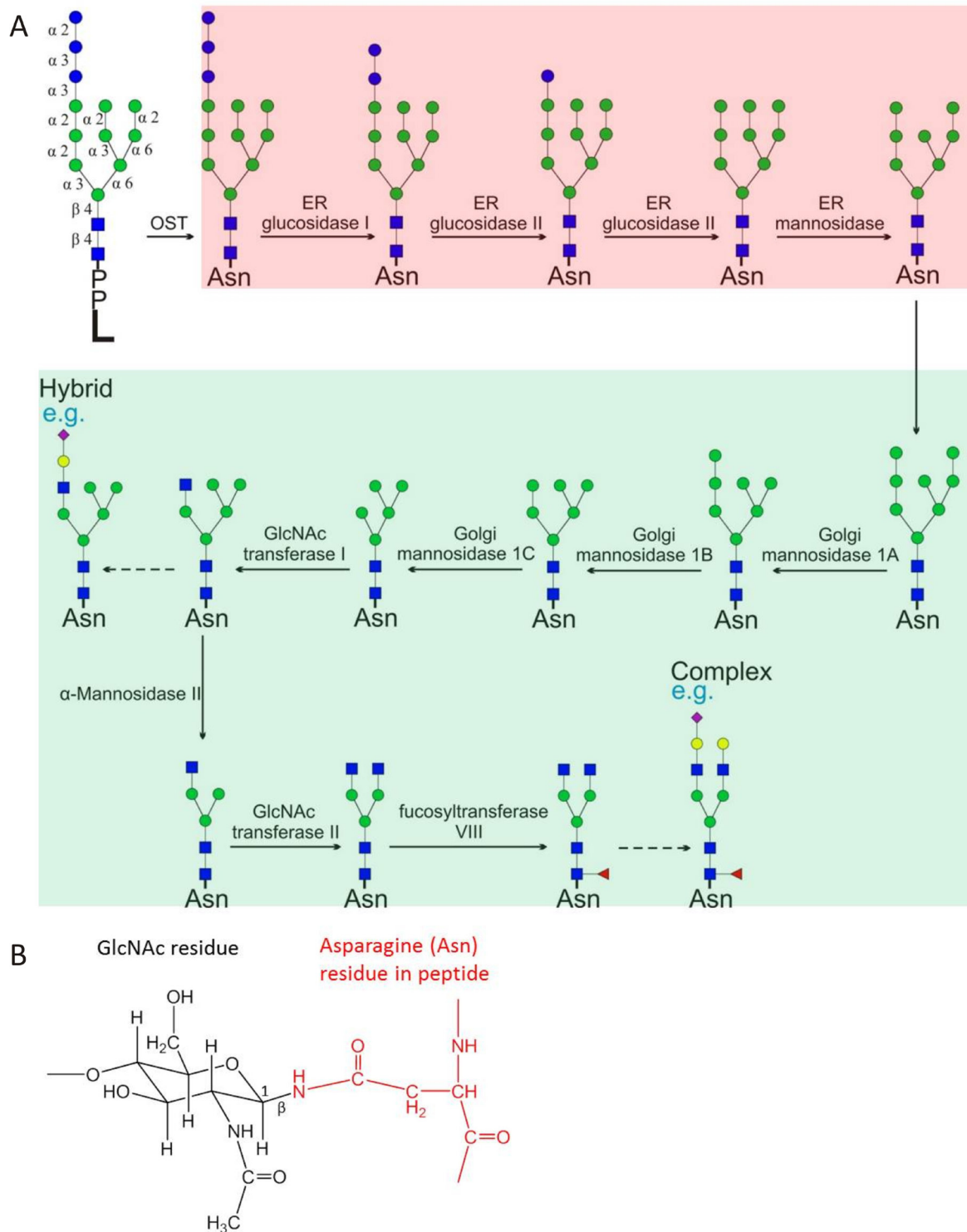
Figure 1.5 A simplified biosynthesis pathway of N-glycans and the structure of a GlcNAc residue on an Asn

A. The linkage in the precursor is labelled. The heptasaccharide (5 Man and 2 GlcNAc) is the core structure of both hybrid and complex N-glycans. In addition to these processes, others can also occur including biosynthesis of polylactosamine antenna, addition of Fuc, and other modifications; **L**, dolichol; PP, a pyrophosphate; the process highlighted in pink shadow is carried out in the ER and the process highlighted in green shadow is performed in the Golgi apparatus; ■ GlcNAc, ● Man, ● Glc, ○ Gal, ▲ Fuc, ◆ NeuAc, OST, oligosaccharyltransferase. **B**. The structure of a GlcNAc (black) attached to an Asn (red) residue. The linkage between the GlcNAc and the Asn is β-form.

A GlcNAc attached to the 4-position of the β-linked core Man by GlcNAc-T III is considered as a bisecting structure which is not considered as an antenna (Figure 1.6). Addition of this GlcNAc requires the prior action of GlcNAc-T I. Bisecting GlcNAc cannot be further extended. *In vitro*, the presence of this structure prevents the actions of GlcNAc-T II, IV and V, α-mannosidase II and core α 1,6-Fuc-T (Schachter, 1991; Takahashi et al., 2009). However, *in vivo* GlcNAc T-III can act after core fucosylation and antennae initiation, so bisected structures can be core fucosylated as well as multiantennary (Hashii et al., 2009; Klisch et al., 2008).
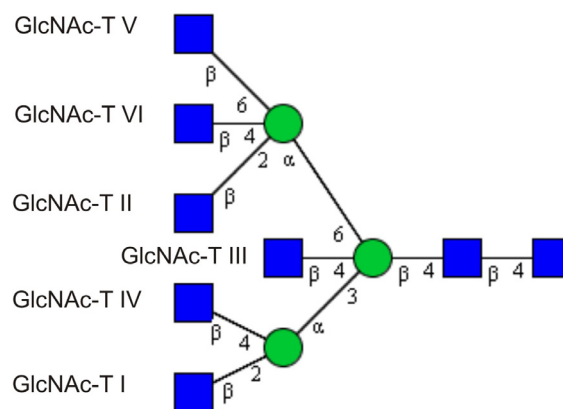
Figure 1.6 Branching of complex N-glycan

The first antenna is initiated by GlcNAc-T I, GlcNAc-T II yields a biantennary glycan, GlcNAc-T III creates a bisecting glycan. More branched complex N-glycans can be produced via the action of GlcNAc-T IV, V and VI. ■ GlcNAc, ● Man.

Complex glycans can have more than four antennae in some vertebrates, but human N-glycans usually possess four antennae at the most. These antennae can be modified with a variety of monosaccharides, such as Fuc and NeuAc. These modifications occur in the *trans*-Golgi apparatus and lead to many terminal structures which can be functionally important. Figure 1.7 shows that an antenna in a biantennary glycan can be modified variously. In addition, the core of this biantennary glycan can also be modified with Fuc, which in mammals, is linked to the 6 position of the proximal GlcNAc. The order of monosaccharide addition and removal is not random but guided, and some enzymes are able to compete for the same substrate, which generates various N-glycans. Nearly all proteins that pass through

the ER–Golgi conduit are N-glycosylated, and a lack of this modification can be fatal in species ranging from yeast to mammals (Freeze, 2006).
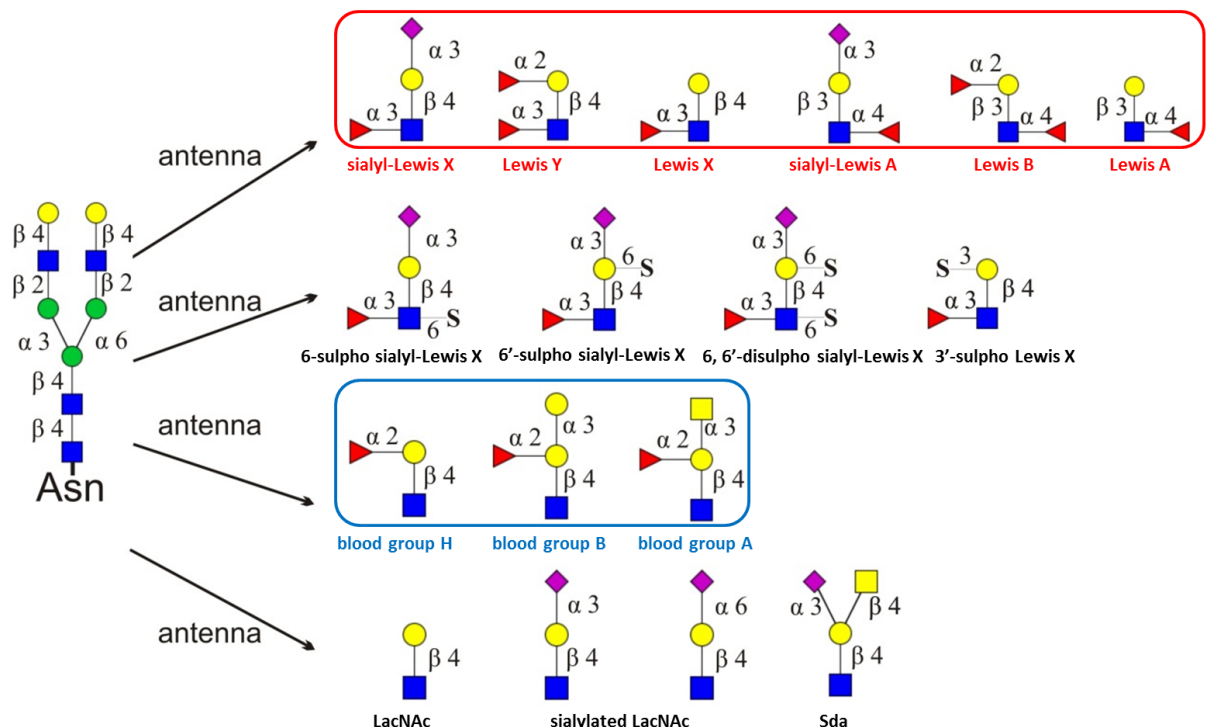
Figure 1.7 Common modifications of the antennae in N-glycans

The antennae labelled in red frame are Lewis blood groups, their names are written in red under the structures. The antennae labelled in blue frame are ABO blood groups, their names are written in blue under the structures. S, sulphate group. ■ GlcNAc, ☐ GalNAc, ○ Gal, ● Man, ▲ Fuc, ◆ NeuAc.

As previously mentioned, each antenna can be extended in the form of LacNAc (N-acetyllactosamine), and then further decorated by Fuc, sialic acid and/or sulpho groups to yield many structures (see Figure 1.7) which could bear variable functional roles. For instance, with L-selectin, optimal recognition involves a sulphated type of sialyl Lewis X, in which the sulphate group is specifically at the 6 position of the GlcNAc residue (Varki, 2007).

Since there are N-glycosylation consensus sites, a program named NetNGlyc 1.0 Server has been developed, which is widely used for looking for the consensus sequence and thus making a general N-glycosylation site prediction (Gupta and Brunak, 2002).

Generally speaking, O-glycosylation is not as well understood as N-linked glycosylation. There are several types of O-glycans that are classified by the first carbohydrate residue which is added to the amino acid residue, such as O-fucosylation for fucose (Fuc), O-glucosylation for glucose (Glc), O-mannosylation for mannose (Man) and O-GalNAcylation for N-acetylgalactosamine (GalNAc) (Wopereis et al., 2006; Martínez-Duncker et al., 2012; Gebauer et al., 2008). Some O-glycosylations can occur in conserved domains, for instance, in epidermal growth factor (EGF)-like domains O-fucosylation and O-glucosylation can occur in conserved sequences Cys-B-Ser-B-Pro-Cys and Cys-B-B-Gly-Gly-Thr/Ser-Cys respectively, in the conserved sequences B can be any amino acid (Harris and Spellman, 1993; Gebauer et al., 2008). However, some other O-glycosylation, such as the O-GalNAc glycosylation, does not occur in conserved sequences. This may be due to the fact that there are various transferases with overlapping but different substrate specificities.

Unlike N-glycans, the biosynthesis of O-glycans is carried out after protein folding in the late ER or in one of the Golgi compartments (Wopereis et al., 2006; Spiro, 2002; Rapoport et al., 1996). The biosynthesis of O-glycans does not start with the *en-bloc* transfer of a dolichol-linked precursor, but with the attachment of a monosaccharide to the target peptide chain (Wopereis et al., 2006).

1.2.2.2.1. O-GalNAc glycosylation

GalNAc-linked O-glycans are often referred to as mucin type (Bergstrom and Xia, 2013). It has been considered as the most differentially and complex controlled type of protein glycosylation (Steentoft et al., 2013; Lira-Navarrete et al., 2015).

O-GalNAc glycosylation is initiated in the Golgi apparatus by the addition of a GalNAc to the oxygen of the hydroxyl group of Ser or Thr forming an O-GalNAc linkage (Figure 1.8), and this reaction is catalysed by polypeptide N-acetylgalactosaminyltransferases (pp-GalNAc-T) (Yoshimura et al., 2012; Ten Hagen et al., 2003). They feature two domains: a catalytic domain and a lectin domain (Lira-Navarrete et al., 2015). It is reported that in humans there are 20 pp-GalNAc-Ts (Bennett et al., 2012; Lira-Navarrete et al., 2015). These

enzymes can be classified into two categories, those that add a GalNAc to an unmodified peptide acceptor and those that favour acceptor peptides containing a GalNAc already (Ten Hagen et al., 2003; Lira-Navarrete et al., 2015). The underlying mechanism of how the GalNAc-T lectin domain contributes to glycopeptide specificity and catalysis is still not clear. Very recently, a group of scientists have presented the first crystal structures of GalNAc-T2 in complex with defined GalNAc-glycopeptide substrates, and they also demonstrated a cooperative mechanism via which the GalNAc-T lectin domain enables unoccupied acceptor site binding of glycopeptides in the catalytic domain (Lira-Navarrete et al., 2015).
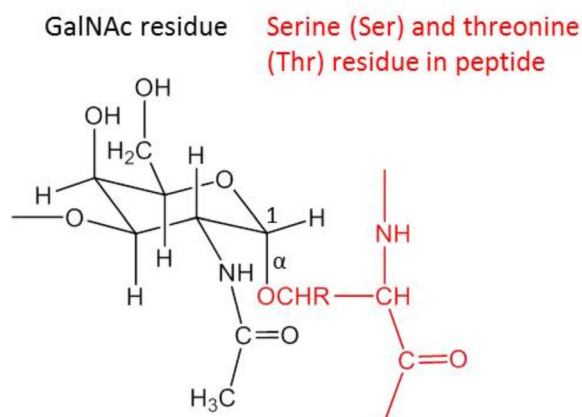
Figure 1.8 The structure of a GalNAc attached to a Ser or a Thr

The linkage between the GalNAc (black) and the Ser or Thr (red) is α-type. R=H in Ser, R = CH_3 in Thr.

O-GalNAc, the simplest mucin O-glycan, is also termed the Tn antigen (Gill et al., 2011). This antigen and its sialylated form (NeuAcα2,6GalNAc) are usually observed in mucins from tumours but not normal tissues or cells (Brockhausen et al., 1995; Julien et al., 2012). Further addition of one or two monosaccharides to the Tn antigen creates 8 core structures of O-glycans (Figure 1.9), 4 of which are most common: core 1 (T antigen), core 2, core 3 and core 4. The T antigen is derived from O-GalNAc by a core 1 β1,3-galactosyltransferase (β1,3Gal-T). In many glycoproteins, the antigen is usually sialylated at the C3 position of the Gal or GlcNAcylated at the C6 position of the GalNAc by a β1,6-N-acetylglucosaminyltransferase creating the core 2 structure. In another pathway the core 3 structure can be created by a core 3 β1,3-N-acetylglucosaminyltransferase. The core 4 structure is yielded by adding a GlcNAc to the core 3 GlcNAc in a β1,6 linkage. The core 5 to core 8 structures show restricted expression: core 5 has been found in human meconium and intestinal adenocarcinoma tissue, core 6 has been detected in human intestinal mucin and

ovarian cyst mucin, core 8 has been reported in human respiratory mucin, and core 7 has only been found in bovine submaxillary mucin (Brockhausen, 1999; Wopereis et al., 2006; Varki, 2009; Chai et al., 1992).

Antennae on O-glycans can be extended in a similar mode to that described in the N-glycan biosynthesis in section 1.2.2.1.
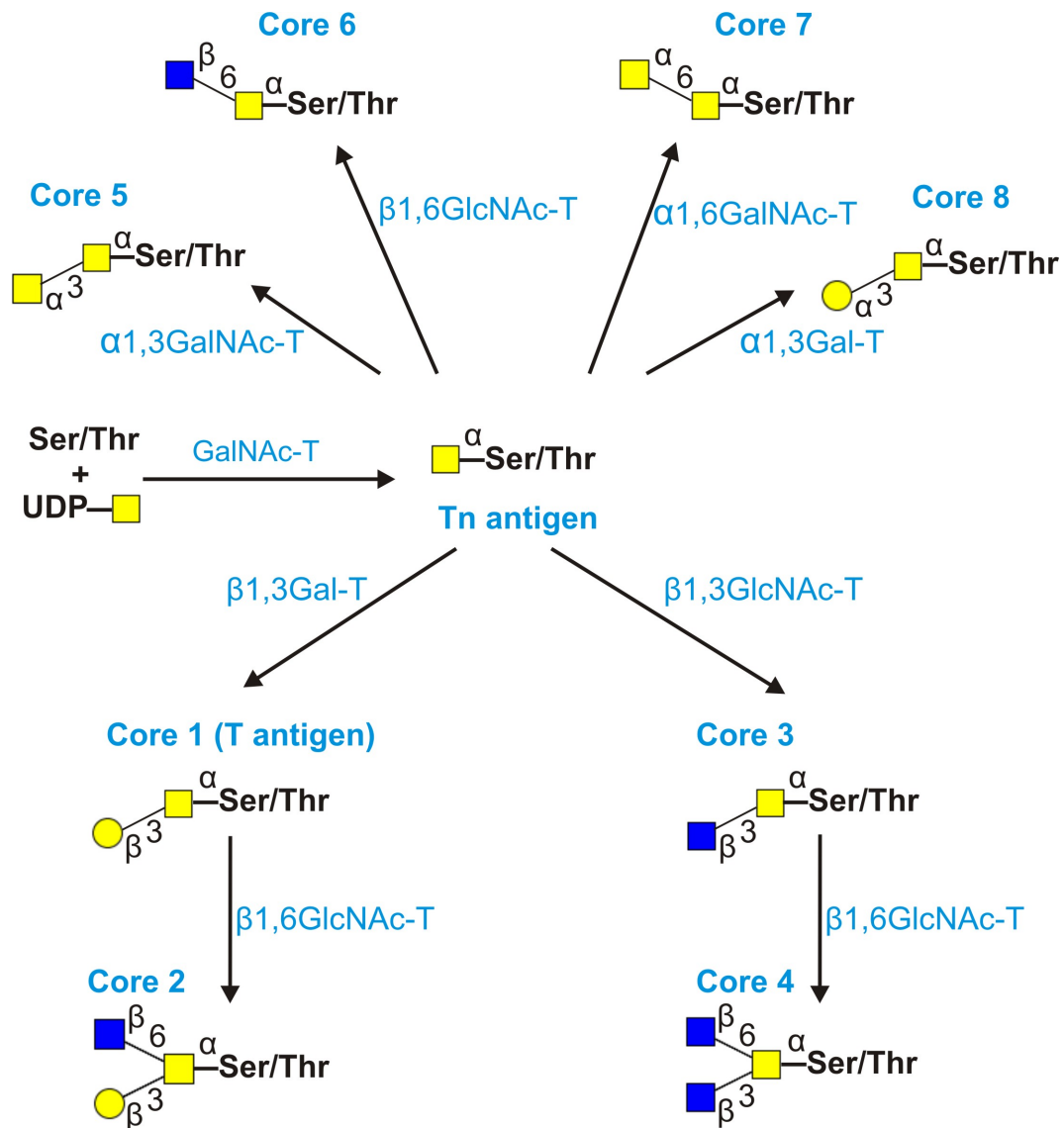
Figure 1.9 Eight Core structures of mucin type O-glycans

The enzymes involved in the synthesis of these core structures are labelled in blue. These structures may be further modified by other glycosyltransferases. ■ GlcNAc, ▢ GalNAc, ◯ Gal.

Although there is no O-GalNAcylation (mucin type) consensus sequence, statistical studies have shown several general rules for the mucin type glycans, such as O-GalNAcylation

preferring sequences rich in Ser, Thr and Pro (Elhammer et al., 1993). Based on these rules, a program called NetOGlyc 4.0 Server has been developed and it is now available on the Internet and is able to predict approximately 75% of mucin-type glycosylated residues correctly (Julenius et al., 2005; Wopereis et al., 2006).

1.2.3. Glycolipid glycosylation

In addition to being attached to proteins, glycans can also be added to lipids, and this type of glycoconjugate is termed glycolipid. Glycolipids play pivotal roles in cell membrane structure, cell-cell interaction, cell-molecule interaction and modulating the function of membrane proteins (Varki, 2009; Lahiri and Futerman, 2007). Based on the locus of the lipid parts, glycolipids can be divided into two categories: glycosphingolipids (GSLs) which are built on a ceramide and glycophospholipids which are located at a phosphatidylglycerol core (Taylor and Drickamer, 2011). GSLs can be found in virtually all plasma membranes of mammalian cells (Yamashita et al., 1999; Kolter and Sandhoff, 1998). In this thesis, the main focus is the glycosylation of GSL.

1.2.3.1.   Glycosphingolipid glycosylation

GSLs are composed of ceramide (Cer), which contains a fatty acid portion and a sphingosine portion and one or more attached saccharides (Figure 1.10). GSL is a functionally important group of glycolipids found in the membranes of cells from bacteria to humans (Jia et al., 2014; Varki, 2009). It functions either like membrane glycoproteins which are involved in cell-cell interactions or in the formation of membrane domains (Taylor and Drickamer, 2011). For instance, one type of glycosphingolipid, gangliosides, can function as toxin receptors (Lahiri and Futerman, 2007).
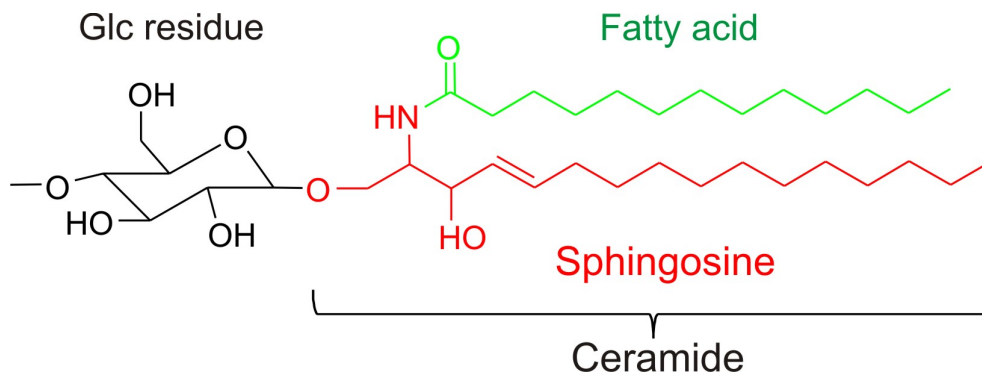
Figure 1.10 The structure of a Glc residue attached to ceramide

The ceramide (Cer) of glycosphingolipids is composed of a long-chain amino alcohol, sphingosine (red), in amide linkage to a fatty acid (green).

Based on the first sugar attached to the sphingosine, GSLs can be divided into two subgroups: when the first sugar is a Gal, the GSL belongs to galactosphingolipid subgroup, when the first sugar is a Glc, the GSL is a member of the glucosphingolipid subgroup (Sillence, 2007). In mammals only Glc and Gal can be added to the Cer directly. GalCer can be sulphated, and its sulphated derivatives are termed sulfatides. The sulfatide is abundant in myelin (Merrill, 2011; Gault et al., 2010). It is reported that the sulfatide plays an important role in the maintenance of CNS myelin and axon structure in mice (Marcus et al., 2006). GlcCer serves as the precursor for complex glycosphingolipids while GalCer is simpler (Merrill, 2011). In this thesis, the glucosphingolipid subgroup is the focus.

The biosynthesis of glucosphingolipids starts with the synthesis of Cer in the ER (D'Angelo et al., 2013; Merrill, 2002; Kolter et al., 2002). The sugar addition is initiated in the Golgi: the Cer moves to the cis-Golgi, probably via vesicular transport (Gault et al., 2010), and the addition of the first Glc residue from a nucleotide sugar donor UDP-Glc to the Cer occurs at the cytosolic surface of the Golgi apparatus. This reaction is catalysed by ceramide glucosyltransferase which was initially detected in an embryonic chicken brain in 1968 (D'Angelo et al., 2013; Funakoshi et al., 2000; Ichikawa et al., 1996; Basu et al., 1968). After this the GlcCer is transferred by a poorly defined GlcCer transporter to trans-Golgi network. This is proposed to be operated by a protein termed four-phosphate adaptor protein 2 (FAPP2) which also controls vesicular trafficking from the Golgi apparatus to the plasma membrane (D'Angelo et al., 2012; D'Angelo et al., 2007). Additional sugars are added from nucleotide sugar donors in the Golgi apparatus. Firstly, lactosylceramide (Galβ1,4GlcβCer) is produced by addition of Gal, catalysed by a β1,4 galactosyltransferase (GalT) (Lannert et al., 1994;

D'Angelo et al., 2013). Further modifications of this structure can yield a series of core structures (Figure 1.11).
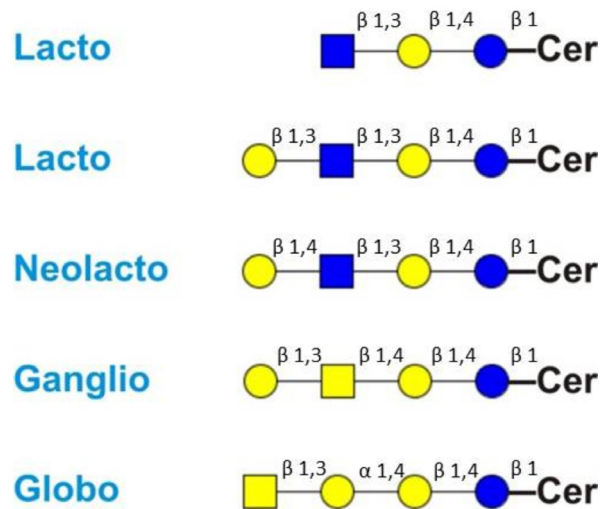
Figure 1.11 Several core structures of glycosphingolipids (GSLs)

GSLs have lacto-, neolacto-, ganglio- and globo-core structures. As shown in the figure, the lacto-series GSLs are based on the core structure Galβ1,3GlcNAcβ1,3Galβ1,4GlcβCer, whereas neolacto-type is built on core structure Galβ1,4GlcNAcβ1,3Galβ1,4GlcβCer, ganglio-series is constructed on Galβ1,3GalNAcβ1,4Galβ1,4GlcβCer, globo-type is formed on GalNAcβ1,3Galα1,4Galβ1,4GlcβCer. ■ GlcNAc, ☐ GalNAc, ● Glc, ○ Gal.

A group of glycosphingolipids were extracted from ganglion cells, and based on this fact this group of glycosphingolipids were called gangliosides. This term was coined by the German biochemist Klenk (Kolter, 2012).

In gangliosides the core is extended by a β1,4 N-acetylgalactosaminyltransferase (GalNAc-T) transferring a GalNAc to the Galβ1,4GlcβCer described previously. In addition, ganglioside biosynthesis involves the activities of Gal and NeuAc transferases, the abbreviations of which are Gal-T and Sia-T shown in Figure 1.12. The ganglioside-glycans are synthesised when they travel through Golgi as the enzymes required are attached at the membrane of the Golgi and function sequentially.
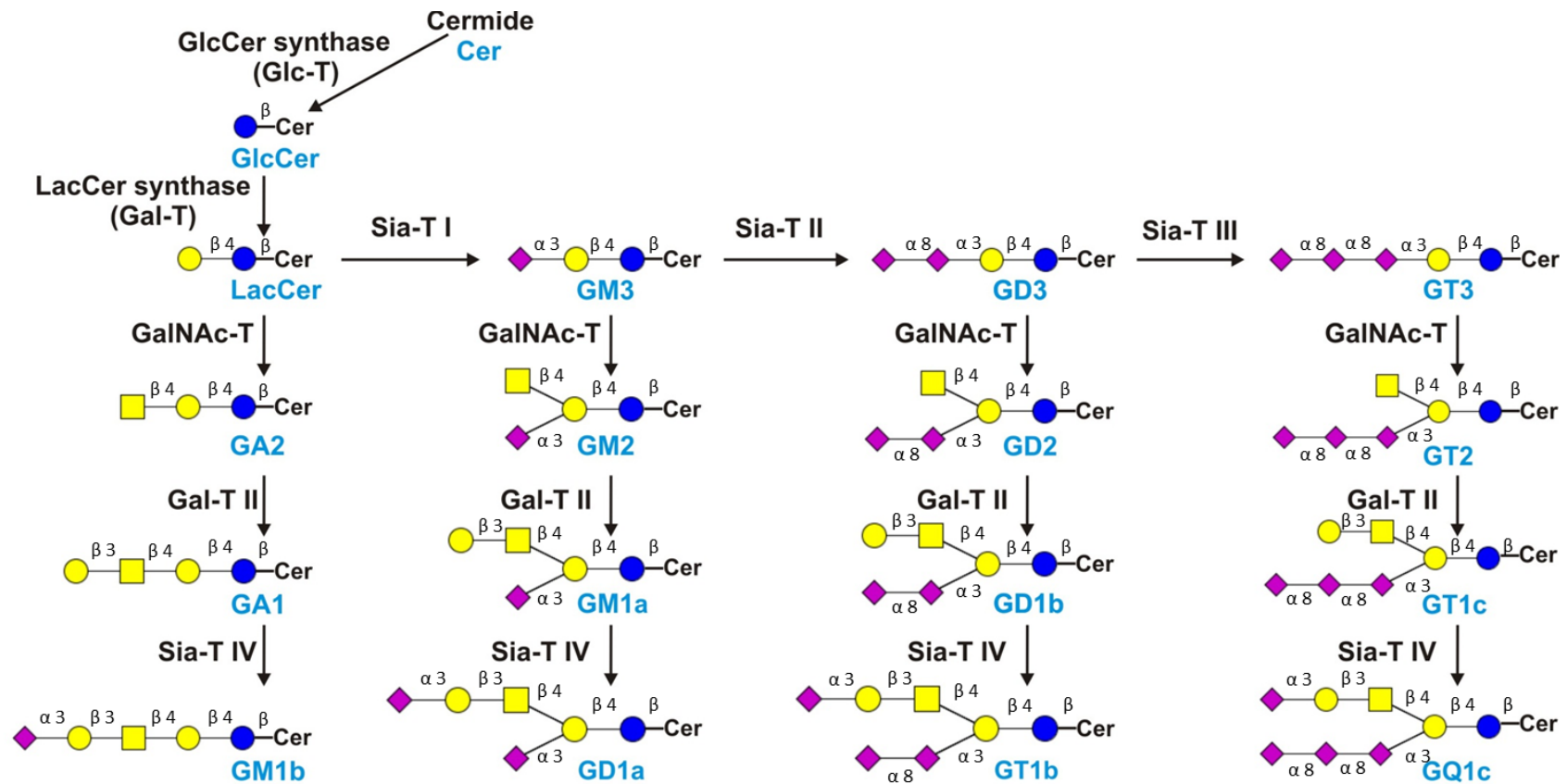
Figure 1.12 Biosynthesis pathway, structure and nomenclature of brain gangliosides

The ganglioside is labelled via Svennerholm nomenclature which is complicated: the G is the abbreviation of ganglioside. The second capital character stands for the sialylation state of every glycans: A means asialylated, M stands for monosialylated, D is short for disialylated, T is the abbreviation of trisialylated and so forth. The number here assigned to individual glycan refers the migration order of this glycan in a certain chromatographic system (Svennerholm, 1963). ☐ GalNAc, ● Glc, ○ Gal, ◆ NeuAc.

As shown in the figure, in some ganglioside-glycans, such as GD1b, GD3 and GT3, there is α2,8 linked NeuAc (Figure 1.12). This linkage has been observed in disialosyl motifs on human O-glycans but not in N-glycans (Traving and Schauer, 1998; Fukuda et al., 1987). This can be explained by an *in vitro* study, the result of which showed that the hST8Sia VI recombinant enzyme has substrate specificity; this enzyme needed the trisaccharide Neu5Acα2,3Galβ1,3GalNAc to produce disialosyl motifs particularly on O-glycans (Teintenier-Lelievre et al., 2005).

1.3. Biological importance of glycans in glycoproteins and GSLs

Since glycoproteins and GSLs are commonly found on cell membranes, they play important roles in cell-cell and cell-molecule interactions.

It has long been realized that proteins with various glycosylation have varied biological properties. Although the glycan portion is considered as a decoration on the proteins, some glycans are also important (Bateman et al., 2010; Yoshida-Moriguchi et al., 2010). For instance, the binding of human sperm to the zona pellucida (ZP) which is the extracellular matrix coating of the oocyte is regulated by the oligosaccharide sialyl-Lewis X [NeuAcα2,3Galβ1,4(Fucα1,3)GlcNAc] on the ovum surface but not Lewis X [Galβ1,4(Fucα1,3)GlcNAc] (Pang et al., 2011).

GSLs are considered to be integral for many vital cell membrane biological events, including cellular interactions, signalling, and trafficking (Yamashita et al., 1999). GSLs are capable of forming microdomains on membranes, and these microdomains are believed to be involved in recognition events and cell signalling (Hakomori, 2004; Gupta and Surolia, 2010). It is known that some growth factor receptors, such as the epidermal growth factor receptor and the nerve growth factor receptor are located in membrane microdomains and there is evidence that signalling functions are considerably regulated by glycolipids (Hakomori, 2003). It has also been reported that in the mouse, in the absence of GSL synthesis, embryogenesis proceeded well into gastrulation with differentiation into primitive germ layers and patterning of the embryo but was finally ceased by an apoptotic process (Yamashita et al., 1999).

The following table gives a few more examples to show the biological importance of glycans.

Table 1.3. A few more examples showing that N-, O- and GSL glycans are biologically important

Glycan type	Related enzyme/gene	Results
N-glycan	Phosphomannose isomerase	Mutations in this gene could cause deficient N-glycan biosynthesis, which could result in congenital disorder of glycosylation type Ib (Lonlay and Seta, 2009).
	Fuc-T VIII	Fut8-knockout mice displayed severe growth retardation, and 70% of the mice died within a few days after birth (Zhao et al., 2008).
O-glycan	Core 1 β1,3Gal-T	Genetic ablation of this enzyme in mice results in defective angiogenesis and fatal brain hemorrhages (Guzman-Aranguez and Argueso, 2010).
	Core 2 GlcNAc-T I	Mutations in this gene lead to defective leukocyte P-selectin ligand activity (Ellies et al., 1998).
GSL glycan	GM2 (see Figure 1.12 for the GM2 structure) synthase	Mice display severe neurological pathology and die soon after birth once the GM2 synthase is deleted (Lingwood, 2011).

Furthermore, the 2012 National Academy Sciences report has demonstrated that a better understanding of glycoscience is essential for advancement of human health and sustainability on the earth (Hart, 2013).

Even the 2014 FDA Guidance for Industry implies the importance of glycans (FDA, 2014): "For proteins that are normally glycosylated, use of a cell substrate production system and appropriate manufacturing methods that glycosylate the therapeutic protein product in a nonimmunogenic manner is recommended ".

All these indicate that the glycosylation and glycans play vital roles in biological processes.

1.4.   Glycomics

Since the importance of glycans has been highlighted in the previous section, it is now necessary to introduce glycomics.

1.4.1. Introduction to glycomics

Glycomics is the comprehensive study of all glycan structures in a provided organism, tissue or cell cluster (Varki, 2009). Glycomic experimental approaches had been employed in all the projects. The term 'glycomics' is composed of two parts: the prefix 'glyco-' which means sweetness, and the suffix"-mics" which relates to a field of study in biology for example genomics and proteomics. Like genomics and proteomics are focused on genes and proteins, respectively, glycomics seeks to illustrate the importance of glycans in biological systems (Cummings and Pierce, 2014). It is believed that the increased understanding of glycan functions is beneficial for drug discovery and disease treatments as there are more than 100 human genetic disorders which are related to deficiencies in the different glycosylation pathways (Freeze et al., 2015). Additionally, some scientists are now paying attention to how carbohydrates affect neuromuscular junction disorders and immune tolerance in the human female reproductive tract (Senderek et al., 2011; Clark and Schust, 2013; Cummings and Pierce, 2014). New findings in these promising areas will be a significant step in biomedical research.

Based on the 'omics' research trend from DNA to proteins of the past decade, it is expected that the next ten years will see a move towards greater exploitation of glycomics.

1.4.2. Difficulties in a glycomic study

Glycosylation is a complicated process. A single protein can be glycosylated heterogeneously. Here the 'heterogeneously' has two meanings, it either means that a glycosylation site is occupied by different glycans or a protein has more than one glycosylation site, but not all of the sites are occupied. For instance, there are two N-glycosylation sites in human transferrin, the locations of which are Asn-432 and Asn-630 respectively (Spik et al., 1975). In normal healthy people, these two sites are primarily occupied by two disialylated N-glycans, however in congenital disorders of glycosylation (CDG) type I patients only one site is primarily glycosylated, and in CDG type II patients two sites are fully occupied by underglycosylated (immature) structures (Freeze et al., 2015; Freeze, 2006). In addition,

unlike DNA and proteins which only have linear structures, glycans usually have antennae, which also complicate the research.

Although traditional biological approaches such as gene knock out and Western blot are able to analyse glycan-related genes and the products, certain glycan-related information may be not clear. For instance, the glycan structures cannot be determined and it is not clear whether the glycosylation sites are fully occupied. In addition, it is also difficult to know the glycosyltransferases and glycosidases which are responsible for transferring and cleaving monosaccharides during glycan biosynthesis. Therefore glycomic research requires continuous development of new techniques which can overcome these difficulties. One of these techniques is mass spectrometry.

1.5. Mass spectrometry

Mass spectrometry (MS) is a technique that measures the mass-to-charge ratios of ions (Calvete, 2014). A mass spectrometer can be used to determine masses of molecules, analyse chemical compounds, and profile the structure of a molecule. MS has played an important role in the "Omics" Era (Di Girolamo et al., 2013). The apparatus is composed of three parts: an ion source, a mass analyser and a detector (Calvete, 2014).

Although MS has been demonstrated since World War I, it could only be used to analyse small molecules due to the undeveloped technology (Keith, 1999). In 1960s, electron ionization was already used for the structural elucidation of di- and trisaccharides (Kochetkov et al., 1968). At that time it was difficult to analyse more complex oligosaccharides as the analyte should be sufficiently volatile and increased molecular weight of the analyte resulted in decreasing volatility. In addition, the mass range of the mass spectrometer also limited the analysis of the oligosaccharides with higher molecular weight. In the early 1970s, the Forssman glycolipid hapten of horse kidney was characterized using mass spectrometry (Karlsson et al., 1974). In this MS analysis, the signals of the molecular ions in the high mass range were generally small or absent, and thus its usefulness was limited. In late 1970s, mass spectrometry was firstly employed in glycoprotein sequencing (Morris et al., 1978; Dell and Morris, 2001). In addition, Dell and Morris carried out the first structural determination of glycans using Fast-atom-bombardment mass-spectrometry (Dell et al., 1983a). These indicate

that MS can be used in glycan related analysis. In the 1980s, the molecular weight of ions that mass spectrometers can deal with did not exceed 10,000 (Keith, 1999). Nowadays, with the development of related technology, MS hardware has been significantly improved. For instance, the employing of matrix-assisted laser desorption ionization tandem time-of-flight (MALDI-TOF/TOF) instrumentation permits the analysis of carbohydrates and proteins with masses up to hundreds of thousands (North et al., 2009). It is therefore considered to be a potent tool for expeditiously and precisely investigating high-mass and fragile molecules.

1.5.1. Ion source techniques

All mass spectrometry experiments require the analyte to be ionized before being analysed in the mass spectrometer. The ionization is achieved either via electron ejection or capture, or protonation/cationization or deprotonation. When choosing an ionization technique, the physicochemical properties of the analyte need to be considered.

In the early days, electron impact or electron ionization (EI) and chemical ionisation (CI) were used (Munson, 1966; Rudge, 1968). However, these techniques are not so suitable for high molecular weight and non-volatile molecules. Since then new techniques, such as fast-atom-bombardment (FAB) and matrix-assisted laser desorption ionization (MALDI), were developed, making direct ionisation and desorption of non-volatile substances possible. Here three ionization techniques EI, FAB and MALDI are introduced in detail as they are either closely related to my glycomic technologies or directly employed in my experiments.

1.5.1.1. *Electron ionization (EI)*

Electron ionization (EI), also known as electron impact, is an ionization technique in which accelerated electrons collide with volatile atoms or molecules to produce ions (Wittmann, 2007; Horning et al., 1977). For ionizing small molecules, EI is widely used due to its advantages: it is a well-understood technique (Rudge, 1968); the ion source is inexpensive compared to others; it is easy to construct as basically it requires a small gap between two electrodes. This ion source is usually connected with gas chromatography (GC).

EI requires the analyte to have volatile and thermally stable properties (Dell and Morris, 2001). During the ionization process (Figure 1.13) a beam of high-energy electrons strikes the analyte and knocks one electron out of the electron cloud of the analyte. Due to the loss of the electron, a singly charged cation $M^{+\cdot}$ is formed. This ion has obtained a large amount of energy from the strike, as a result of which it fragments subsequently, yielding fragment ions with various relative abundances, which provides a 'fingerprint' for the analyte structure (Wittmann, 2007; Busch, 1995).
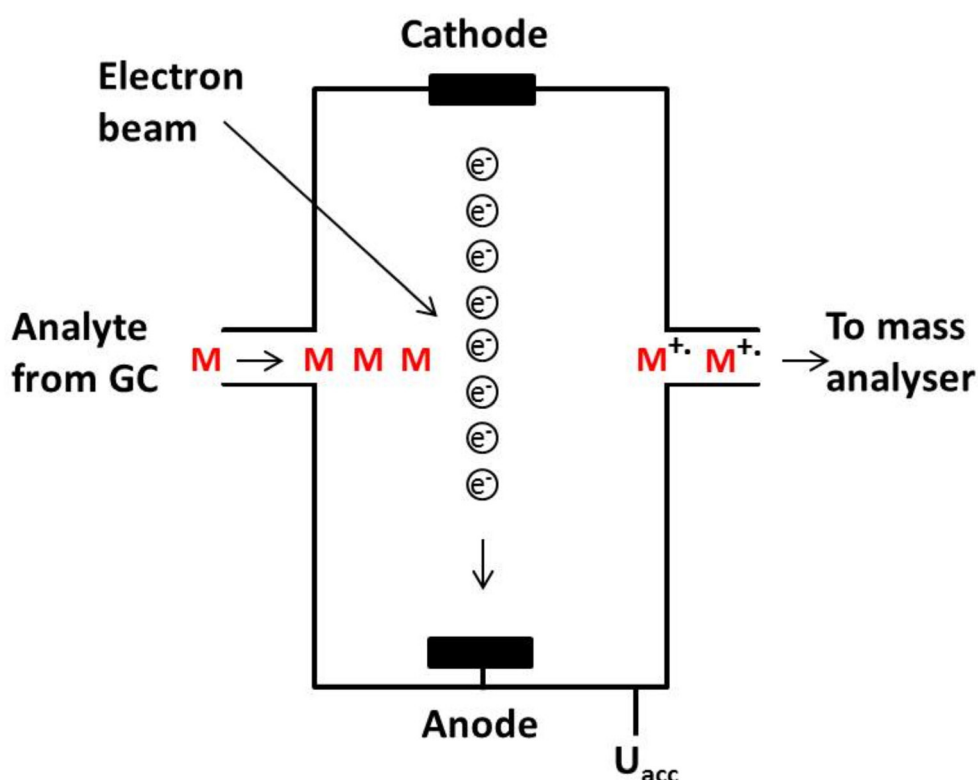
Figure 1.13 Schematic view of the EI ion source

The red M stands for an analyte molecule; U_{acc} is the electron accelerating potential, GC is short for gas chromatography. The efficiency of this technique and the yield of fragments are mainly determined by the property of the analyte and the energy of the electron beam. If the energy of the electrons is much lower than required, the interaction between the electron and the analyte is not able to transfer adequate energy, as a result of which the ionization will not occur. If the energy of the electron is much higher than the required, the ionization will not occur either. Only when the energy is around the required value does the De Broglie wavelength of the electron match the length of the chemical bonds in the analyte, and maximum energy can transfer to the analyte, resulting in ionization.

Although this technique works well for volatile analytes, it causes substantial fragmentation, as a result of which the molecular ions are usually absent or very weak. Thus EI is considered as a 'hard' (it causes the ion to fragment) ionization technique (Hejazi et al., 2009; Horning et al., 1977; Dell and Morris, 1974; Sutton-Smith and Dell, 2006).

1.5.1.2.   *Fast atom bombardment (FAB)*

FAB has become a potent structural analysis tool since the reports of its application in 1981 (Dell et al., 1981; Barber et al., 1981a; Barber et al., 1981d; Morris et al., 1981b). This ionization technique is usually coupled with a magnetic sector mass analyser. It has been employed to analyse inorganic ion clusters to mass 25800 Da and peptides to mass 5700 Da in the early 1980s (Rinehart, 1982). In FAB (Figure 1.14), the sample to be analysed was mixed with a viscous and non-volatile matrix (usually thioglycerol) and bombarded in vacuum by a high energy beam of ions or atoms (Rinehart, 1982; Barber et al., 1981c; Dell and Morris, 2001). The bombardment transfers energy to the sample molecule in the matrix, which allows the molecule to sputter out of the matrix and into the vacuum of the ion source (Dell and Morris, 2001). In this process the majority of the molecules couple with a single ion such as H^+ and produce a singly charged cation such as $[M+H]^+$ (Barber et al., 1981c; Barber et al., 1981b; Dell and Morris, 2001). The aim of using a matrix is to minimise sample damage by the atom or ion beam and to absorb spare energy from the beam and thus to prevent extensive fragmentation of the sample.

The main advantage of using FAB in glycan mixture analysis is that an individual FAB-MS experiment is able to give both sequence and compositional information which are obtained from fragment ions and molecular ions, respectively. Thus it has been used to determine structures of interest within a complex mixture of glycans such as biologically functional terminal epitopes (Dell and Morris, 2001; Dell, 1987). For instance, the presence of sialyl-Lewis X on leukocytes was first observed using FAB (Fukuda et al., 1984). In addition, compared to EI it is not limited to analysing volatile samples.

When FAB began a revolution in the MS analysis of glycoconjugates, it was soon recognized, via research at Imperial College London in the early 1980s, that permethylation enhanced its versatility (Bern et al., 2013; Oates et al., 1985; North et al., 2009). This procedure remains unaltered from that used nearly 30 years ago, and it is still applicable to MALDI-MS nowadays (North et al., 2010). This is due to the fact that glycan permethylation significantly improves sensitivity of detection, and the fragment ions yielded from the permethylated glycans are more reproducible than those derived from their native counterparts (Sutton-Smith and Dell, 2006; Dell et al., 1983b).

Due to the limited availability of many biological samples, and their high molecular weights, high sensitivity and extended mass ranges became increasingly important. From the end of the 1990s, FAB was replaced by more sensitive techniques, notably MALDI, which were capable of analysing much larger molecules.
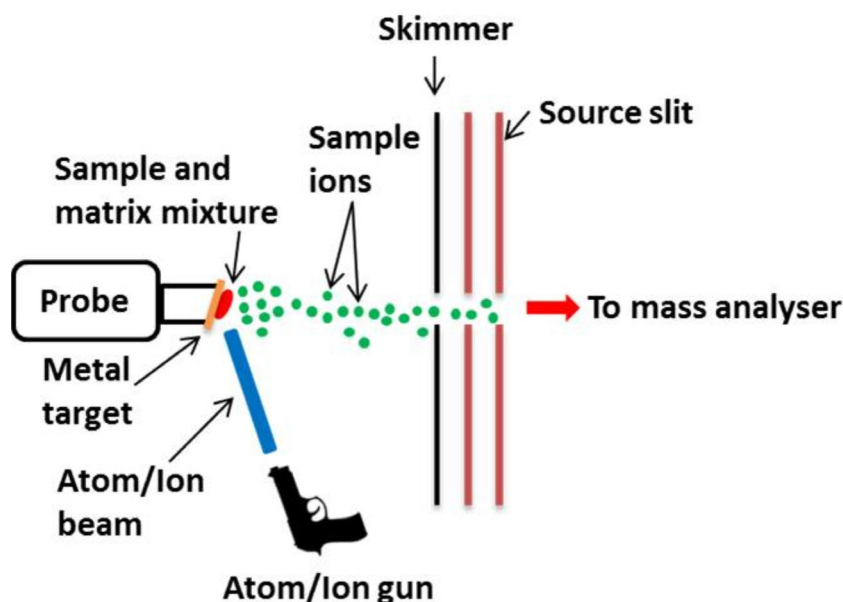
Figure 1.14 Schematic view of the FAB ionization source

1.5.1.3.   *Matrix-assisted laser desorption ionization (MALDI)*

MALDI is currently one of the most sensitive ionization methods for glycan analysis (Loo et al., 1999; Varki, 2009; Kailemia et al., 2014). This ionization technique is usually coupled with time-of-flight (TOF) mass analysers. The wide application of MALDI in the biological area has especially promoted innovation in the design of TOF analysers (see section 1.5.2.2), which results in enhanced mass resolution and sensitivity (Radionova et al., 2015).

In 1963 the laser desorption ionization technique was introduced by Honig and Woolston (Honig and Woolston, 1963) and later improved by Posthumus and employed in biomolecular studies (Posthumus et al., 1978). In the 1980s, this technique was further developed by Hillenkamp and Tanaka (Tanaka et al., 1988; Karas et al., 1985). The 2002 Nobel Prize in chemistry was awarded to the latter due to his contribution to the soft desorption ionisation techniques in MS.

As indicated in the term MALDI, a matrix is involved in the ionization process: the analyte is mixed with a large excess of low molecular weight matrix which can absorb laser energy, and this prevents fragmentation of the sample molecules. The analyte and matrix mixture is loaded onto a metal plate. The analyte co-crystallises with the matrix when the mixture is dried. When the laser is fired at the dried mixture under vacuum conditions, a large amount of energy from the laser is absorbed by the matrix, as a result of which desorption of sample and matrix molecules at the surface layer occurs and a plume forms. Finally, the analyte molecules are ionized (Glish and Vachet, 2003). The velocity that the ions obtained in MALDI is independent of their molecular mass (Beavis and Chait, 1991). High voltage is applied to the metal plate, which accelerates the resulting ions out of the ion source into the mass analyser, which is usually a time-of-flight (TOF) analyser. Unlike EI, the processes of volatilization and ionization in MALDI are intimately related (Busch, 1995). However, the mechanism of the ionization is still poorly understood, but it is widely accepted that the energy from the laser promotes the ion formation by proton transfer in the matrix plume and chemical reactions between the excited matrix and the analyte (Zenobi and Knochenmuss, 1998; Knochenmuss, 2006). Figure 1.15 shows a schematic view of the most widely accepted ion formation mechanism in MALDI.

The analyte molecule ionized by MALDI forms mainly monocharged ions which are denoted as $[M+H]^+$ or $[M+Na]^+$ (Zaia, 2004), thus it is rapid and convenient for the determination of the analyte molecular weight.
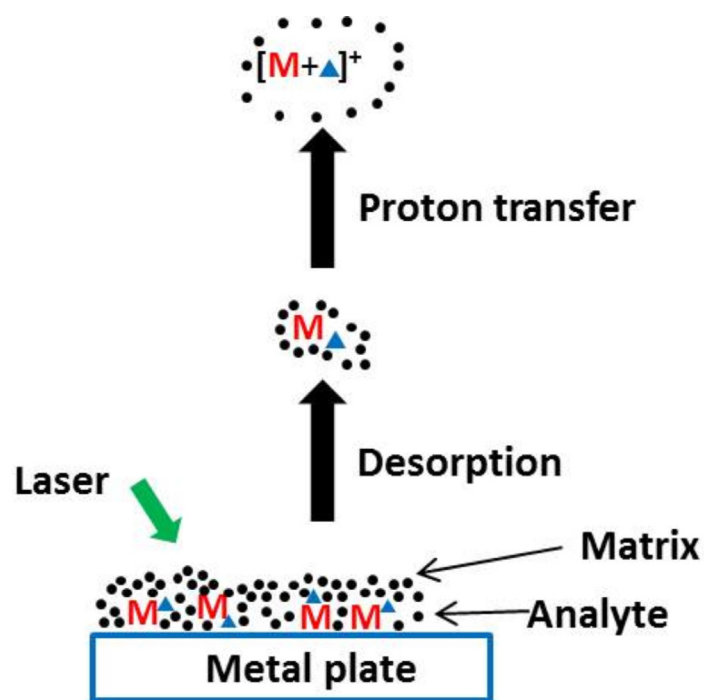
Figure 1.15 Schematic view of the most widely accepted ion formation mechanism in MALDI

The red M stands for an analyte molecule, the black spots are matrix molecules, the blue triangle can be either H or Na.

The resolution of the MALDI mass spectrum can be improved by "delayed extraction" which is basically a time delay between ion generation and ion extraction from the source. The high voltage applied to accelerate the ions is not continuous but pulsed. To some extent this narrows the initial kinetic energy distribution and corrects the energy dispersion of the ions, as a result of which the ions fly with more uniform kinetic energies improving resolution (Kaufmann et al., 1996; Vestal et al., 1995; Vestal and Campbell, 2005; Signor and Erba, 2013).

1.5.2. Mass analysers

The ions need to be separated once they have been produced. A mass analyser is a part of a mass spectrometer that takes control of ionized microscopic particles and separates these particles based on their mass to charge ratios. This form of separation method was originally discovered by Joseph John Thomson who discovered the electron and measured its mass-to-charge ratio and thus won the 1906 Nobel Prize in Physics. Mass analysers employ either

electric or magnetic fields or a combination of both to achieve the separation. Three types of mass analyser will be introduced in detail as they are either more commonly coupled with the previously mentioned ionization techniques or utilised in my projects.

1.5.2.1. *Magnetic sector*

Magnetic fields have been employed in MS since A.J. Dempster developed the first spectrometer with a sector-shaped magnet in 1918 (Münzenberg, 2013). The development of high-field magnets in the 1970s, pioneered by Morris (Morris et al., 1981a), had a profound effect on biological MS. Thus, employing a combination of FAB and high field magnet technology allowed MS of analytes whose molecular weight could go up to 6000 Dalton (Da) (Dell and Morris, 1982). Approximately 15 years ago researchers in our laboratory still used a magnetic sector mass analyser coupled with FAB. Although magnetic mass analysers are expensive, they are still popular in certain type of work, for example, it can be employed in environmental sciences for the analysis of dioxins and furans (Hernandez et al., 2012) as well as mass separation of rare isotopes at accelerator facilities (Münzenberg, 2013).

The passage of analyte ions through a magnetic sector analyser is as follows: the ions generated from the ion source are accelerated via a relatively high potential (Figure 1.16), usually the accelerating potential U is 7-8 keV. The kinetic energy (E_k) that an ion gains is equal to the potential loss, which means that $E_k=mv^2/2=Uq$ (m is the mass of the ion, q stands for the charge of the ion, v is the velocity of the ion). Once the ions pass this acceleration region, they come into a magnetic field, in which they will be separated according to their mass to charge ratios (m/z). In this magnetic field there is a flight tube with a bend of fixed radius r permitting only the ions possessing a required mass-to-charge ratio to travel through. Once the ions enter the magnetic field, they obtain a magnetic force $F_B=qvB\sin\theta$ (because the B is applied perpendicular to the velocity direction, $\sin\theta=\sin90^\circ=1$). This magnetic force make the ions fly in a circular trajectory with the radius r, thus F_B is equal to the centrifugal force, which means $qvB=mv^2/r=2Uq/r$. Finally the equation can be written as $m/q=r^2B^2/2U$ (m/q is the mass to charge ratio m/z), which provides the theoretical basis for separating ions according to their m/z. Only the ion with selected m/z can be transmitted to the detector, those ions with higher or lower m/z will be deflected. By scanning the magnetic field

($m/q=r^2B^2/2U$, fixing r and altering either U or B) ions with different m/z ions are detected, as a result of which a mass spectrum can be obtained (Glish and Vachet, 2003).
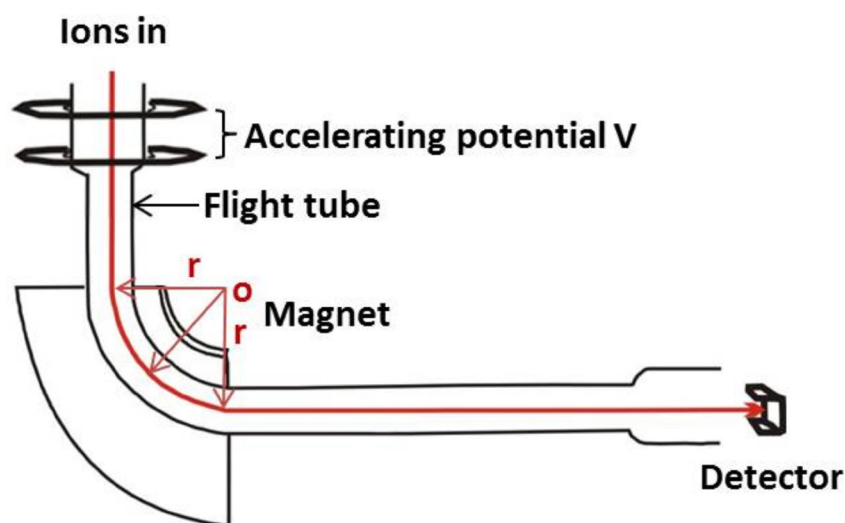
Figure 1.16 Schematic of the magnetic sector analyser

However, due to the initial kinetic energy (E_k) difference the analyte ions with the same m/z can be dispersed by the magnetic field as $r=(2mE_k)^{1/2}/qB$. In order to overcome this, the E_k difference needs to be controlled. To achieve this, an electrostatic sector can be incorporated with the magnetic sector (Nier, 1955). The electrostatic sector functions as an energy focusing device, which is able to increase the resolution of the instrument. This two-sector instrument is called double focusing as in this instrument both the energies and the angular dispersions are focused (Münzenberg, 2013).

1.5.2.2. *Time-of-flight (TOF)*

The concept of a time-of-flight mass analyser was initially mentioned by William in 1946 at MIT (Stephens, 1946). This mass analyser separates ions with different m/z by the flight time difference due to their various velocities in a flight tube without the use of an external field (Cameron and Eggers, 1948; Guilhaus, 1995).

The travelling of ions via a TOF mass analyser can be described as follows. Analyte ions are produced in the ion source. These ions are accelerated via a known potential U. In this

electric field, the electric potential energy of one ion will be converted into its kinetic energy, $Uq = E_k = mv^2/2$ (m is the mass of the ion, v is the velocity of the ion, U is the accelerating potential, q stands for the charge of the ion). After the acceleration, this ion will travel linearly in the flight tube at constant velocity to the detector. The time that the ion requires to travel is $t = L/v = L(m/q)^{1/2}(1/2U)^{1/2}$ (L is the length of tube). As shown in the equation, t, time of flight is directly proportional to $(m/q)^{1/2}$ (L and U are held constant, m/q is m/z) (Guilhaus, 1995). If all other factors are equal, the equation also shows that, the smaller the mass of an ion, the shorter the time it requires to reach the detector.

Theoretically, this type of mass analyser does not have upper mass limit. In 1996 a sample with mass over 333 kDa was measured using MALDI-TOF (Moniatte et al., 1996).

Historically poor mass resolution was a drawback of the TOF analyser. The resolution is mainly limited by the existence of a spread in the initial velocities of analyte ions formed in the source (Mamyrin et al., 1973). Although velocity filters were used to increase the resolution, they also reduced the sensitivity. With the development of numerous techniques, such as the application of a reflectron, the resolution of the TOF mass analyser has been improved.

The reflectron was initially proposed by Mamyrin in 1973 (Mamyrin et al., 1973). The principle of a reflectron is that a retarding field created by the reflectron functions as an ion mirror to deflect the ions penetrating the field and send them back to reach the detector. It corrects the kinetic energy of dispersion of the analyte ions; both fast and slow ions with the same m/z reach the detector simultaneously rather than at different time, which narrows the peak width of the output signal, thus improving the resolution (Figure 1.17).
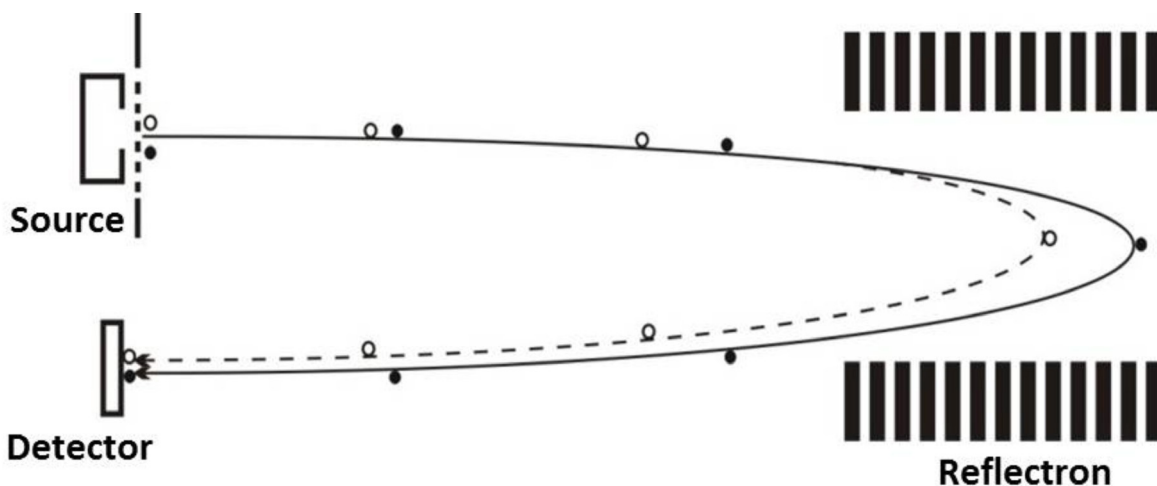
Figure 1.17 Schematic of a TOF mass spectrometer equipped with a reflectron

○ represents the ion with greater energy which travels deeper in the reflectron and follows the full line trajectory; ● the ion with smaller energy travels shallower in the reflectron and follows the dotted line trajectory. These two ions arrive at the detector simultaneously via the adjusting voltages and fields.

The resolution can also be improved via decreasing the accelerating potential. However, decreasing the potential is not workable as it leads to lower sensitivity.

It is widely acknowledged that the TOF mass spectrometer has a number of advantages over other mass spectrometers: the ability of obtaining a broad mass range spectrum in microseconds; the theoretically unlimited mass range for the ions analysed and the possibility of displaying both the full mass spectrum and the individual section of interest (Mamyrin et al., 1973).

TOF is usually coupled with MALDI. A main reason for this is that MALDI delayed extraction (see section 1.5.1.3) is ideally suitable for TOF analysis. The following MALDI-TOF mass spectrometer, Voyager-DE^{TM} STR (Figure 1.18), is the instrument mainly employed in my projects to collect the MS data. When the instrument operates in linear mode, the linear detector collects the ions. However, when the instrument runs in reflectron mode, the reflectron is switched on and it deflects the ions to the reflected detector. The linear mode is usually employed in intact protein studies as it does not lose efficiency for large molecules while the reflectron mode does.
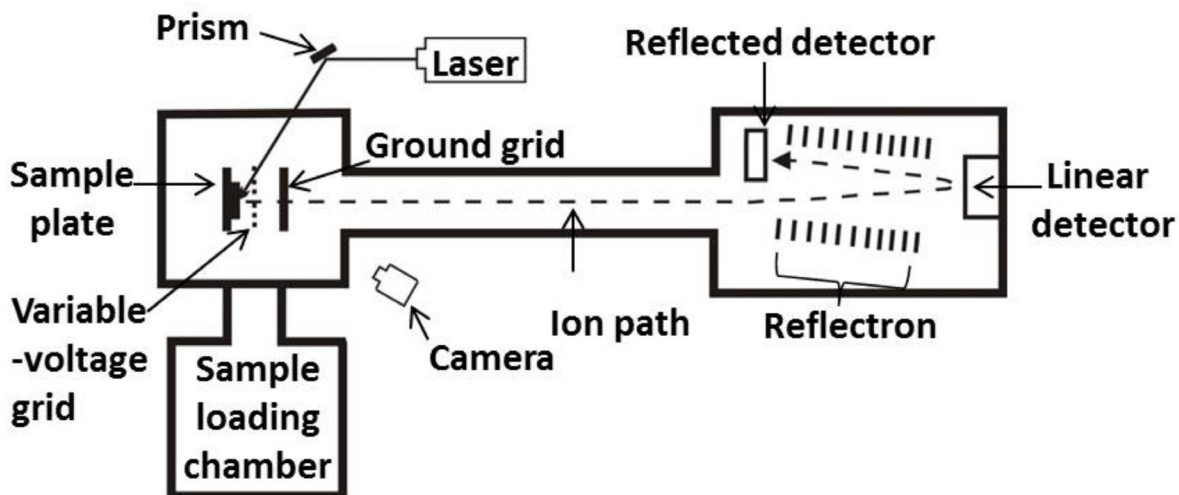
Figure 1.18 Schematic of the Voyager-DE^{TM} STR

Delayed extraction (variable-voltage grid and ground grid) and a reflectron are applied in this instrument. Practically the reflection angle is less than 2°, which makes the displacement along the y axis smaller. In the above figure, a much greater angle has been used for clarity in the drawings. The TOF analyser with a reflectron allows the ions to travel longer, which is able to improve the resolution of ions with similar masses as the longer the ions travel the greater their arrival time difference will be.

1.5.2.3. Quadrupole (Q)

This type of mass analyser was described by Paul and Steinwedel in 1953 at Bonn University, and later they patented this apparatus (Paul and Steinwedel, 1960; Keith, 1999).

As the term implies this mass analyser is made up of four parallel metal rods which are circular or ideally, hyperbolic. It is an instrument which sorts ions according to their m/z ratios using the stability of the trajectories in oscillating electric fields (Paul and Steinwedel, 1960). As shown in Figure 1.19, one rod is paired with the diagonally opposite rod, and thus there are two pairs in total. Each pair is electrically connected; a direct current (DC) potential is imposed positively on one pair of rods and negatively on the other, then a radio frequency (RF) voltage is superimposed on the DC potential, as a result of which one pair of rods is positively charged and the other is negatively charged. These potentials generate a total electric field which has an influence on ions travelling in the quadrupole between the metal rods. Since there is an RF voltage, the generated electric field varies as time goes by and it only allows an ion with a stable trajectory to travel through the quadrupole (Figure 1.19). For instance, when a positive ion enters the analyser, it undergoes an electric field force which

pulls it to a negative rod. If the potential changes sign, the ion will change direction and not hit the rod. The path of the ion will be stable if it never reaches the rod. Ion oscillations depend on the total electric field; only an ion with a specific m/z travelling in a stable trajectory is able to pass through the analyser (Glish and Vachet, 2003).

A mass spectrum is built up via the scan of RF taking ions of successively higher to lower m/z values into focus on the detector. Consequently, the rods function as a mass filter. The quadrupole mass analyser has a number of advantages, including robustness and low cost (Wittmann, 2007). It is compatible with the EI source, and is the most common analyser in GC-MS (Glish and Vachet, 2003).
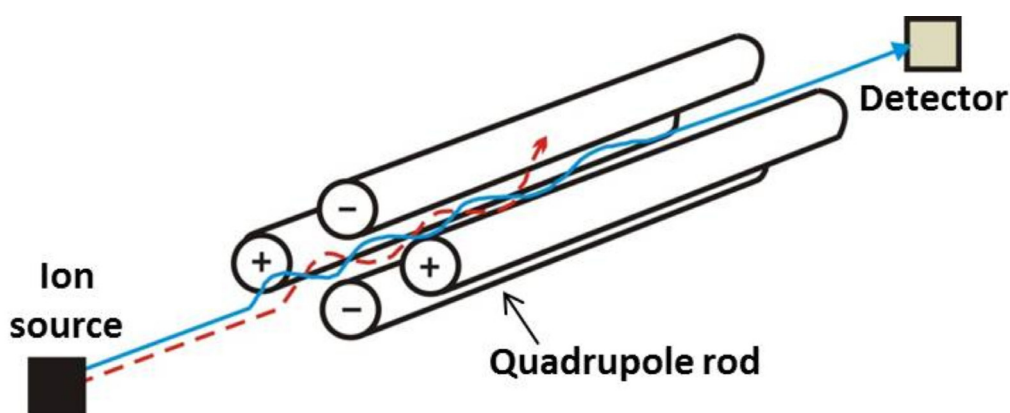
Figure 1.19 The path of ions travelling through a quadrupole mass analyser

Two ions are generated from the ion source. The blue line is a stable trajectory of an ion passing through the analyser, the red dotted line stands for an unstable trajectory of an ion which is intercepted by a rod and thus this ion cannot reach the detector.

1.5.3. The detector

The detector is the final element of a mass spectrometer. Once the ions pass through the mass analyser and then reach the detector, they will be converted into usable signals. The detector either provides a direct measurement of ion charges or amplifies the signal. There are two types of detectors: some detectors are manufactured to count ions with a single mass at a time and thus they count all the ions in sequence at one point (point detectors); other detectors, such as photographic plates and image current detectors, are designed to count ions with different masses at the same time and detect the arrival of these ions in the same plane (array detectors) (Barnes and Hieftje, 2004).

The microchannel plate electron multiplier, a type of array detector, consists of a group of parallel and small channels which are covered with a secondary electron emitting semiconductor substance and drilled into a plate. Electron multiplication is achieved via this substance giving off secondary electrons. The characteristic of this detector is the short response time resulting from the short path of the secondary electron inside the channel, and this is well suited to the TOF analyser which requires precise arrival time. However, this detector has some disadvantages; it is fragile, sensitive to air and expensive.

1.5.4. Tandem mass spectrometry (MS/MS)

As mentioned in section 1.5.1.3, MALDI is a soft ionization technique, producing very limited fragment ions. Therefore, it is suitable for glycan profiling. However, to further analyse monosaccharide composition and sequence of each glycan, fragmentation of the glycan ion is required. The fragmentation allows rigorous sequencing which usually includes glycan branch determination and terminal epitope confirmation. This fragmentation can be achieved by employing two mass analysers in tandem. The term MS/MS refers to this type of coupling.

In this technique, there are two mass analysers which are separated by a collision cell. The first mass analyser functions as a 'filter' to choose a precursor molecular ion of interest from the ion cluster yielded at the source. The chosen ion then enters a collision cell filled with an inert gas such as argon and encounters collision induced dissociation (CID). This collision facilitates fragmentation of the precursor ion; in which the cleavage of glycosidic bonds occurs. The fragments produced in this process continue flying towards the second analyser, which separates these fragments thus generating a fingerprint pattern of the precursor ion. This fingerprint contains sequence informative fragment ions which provide vital structural information. These two mass analysers could be the same or different, for instance they could be tandem time-of-flight (TOF/TOF) analysers or a combination of quadrupole and time-of-flight (Q-TOF) (Morris et al., 1997; Morris et al., 1996b; Vestal and Campbell, 2005).

Tandem MS has now become a potent tool for characterizing components in glycans. All the MS/MS data mentioned in the thesis were obtained using an Applied Biosystems SCIEX

MALDI-TOF/TOF mass spectrometer. Therefore it is necessary to introduce TOF/TOF in detail.

1.5.4.1.   *Tandem time-of-flight (TOF/TOF)*

The schematic of the MALDI-TOF/TOF mass spectrometer, Applied Biosystems SCIEX 4800, used in my Ph.D. study is shown (Figure 1.20). This MALDI-TOF/TOF instrument is composed of a linear delayed extraction MALDI-TOF, a time ion selector (TIS), a collision cell and one more TOF analyser. In the MS mode, ions produced are directly guided to the detector. However, in the MS/MS mode, ions need to pass through the first TOF and then are selected by the TIS. Only ions of interest can pass through. The collision energy is adjustable. The adjustment is achieved by modulating the ion source potential to that of the collision cell, in which the inert gas pressure is managed to make the fragmentation acceptable. The mass spectrometer is able to yield CID fragment ions at both high and low energies, and the resolution and sensitivity of the single TOF can be ensured simultaneously (Vestal and Campbell, 2005). Fragment ions yielded in the collision cell are then re-accelerated towards a reflectron which deflects the ions to a reflected detector.
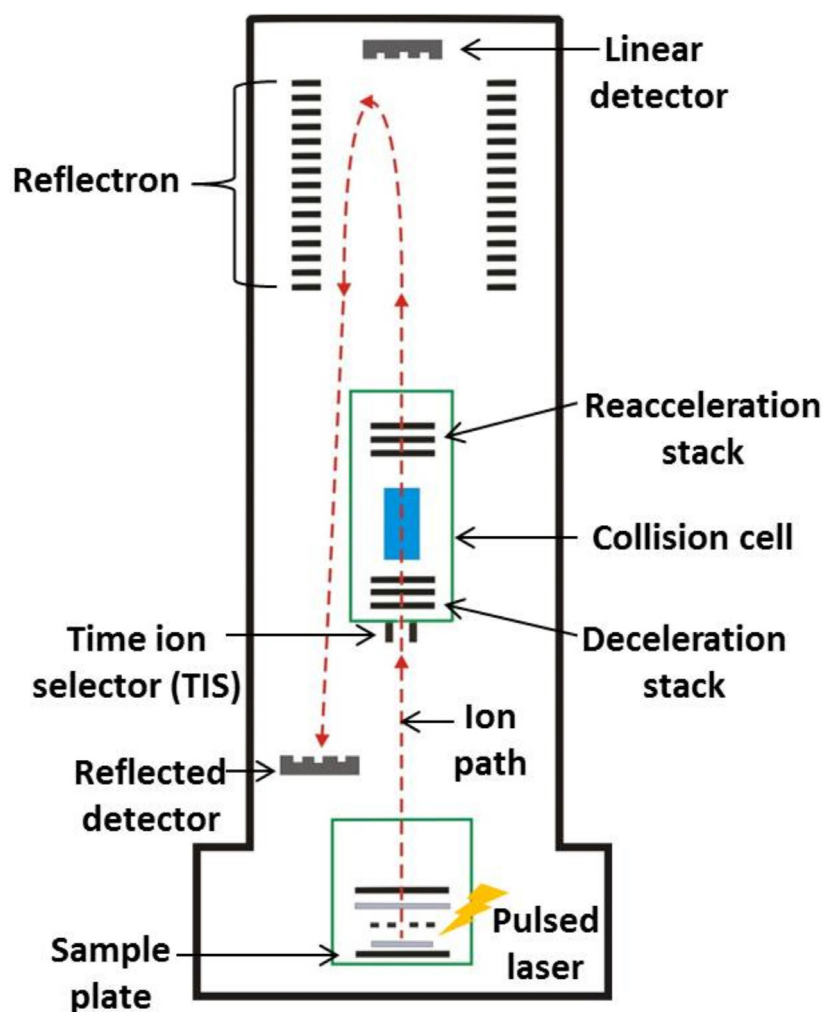
Figure 1.20 A schematic illustration of 4800 MALDI TOF/TOF analyser in MS/MS mode

Ions yielded in the source are flying upwards. The time ion selector (TIS) selects the ion of interest which will be firstly decelerated by the deceleration stack and then fragmented in the collision cell which is blue in the figure.

1.5.4.2.   Fragmentation patterns of permethylated glycans

To interpret the MS/MS spectrum, it is necessary to know the fragmentation patterns of glycans. Currently most of the knowledge concerning glycan fragmentation patterns is originally from the work carried out using FAB (Dell, 1987; Dell et al., 1983a; Egge et al., 1983). All the MS/MS data in this thesis were obtained via MALDI, but carbohydrate fragmentation patterns established by FAB are still valid.

Compared to a native glycan, a permethylated glycan yields a limited number of fragment ions, which makes the data interpretation easier. In addition, its fragmentation is more

reliable (Sutton-Smith and Dell, 2006; Dell, 1990; Yu et al., 2006). Therefore glycans in my projects were permethylated prior to being analysed.

The fragmentation of glycans in MS/MS is mainly due to glycosidic bond breakage between two monosaccharide rings. This breakage can be achieved via a cleavage termed β-cleavage in which a hydrogen transfer is involved. This cleavage does not yield any charge. The charge on the fragment ion is produced via protonation or cationization (e.g. sodiation). The fragment ion can be either reducing or non-reducing, and this depends on which of the two bonds to the middle oxygen is cleaved (Figure 1.21) (Dell et al., 1994; Dell, 1987).
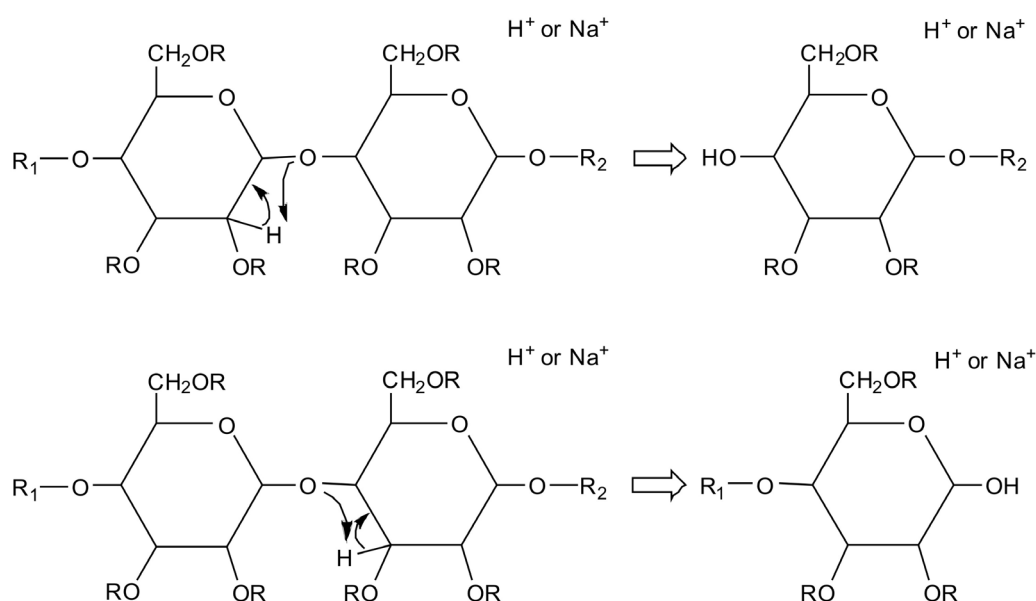
Figure 1.21 β-cleavage occurring at a glycosidic bond can result in either a reducing ion (top panel) or a non-reducing ion (bottom panel)

R_1 is the non-reducing end, R_2 is the reducing end. R stands for the 'remainder' of the chemical group.

In addition to β-cleavage, there is another type of cleavage, A type-cleavage, which also leads to glycan fragmentation (Dell, 1987). The cleavage occurs at the left side of the middle oxygen, it produces an oxonium ion (Figure 1.22). This cleavage favours occurring at the amino sugar residues. In addition, it is sometimes accompanied by a further fragmentation in which the substituent at the 3 position of the ring is eliminated (Dell et al., 1994).
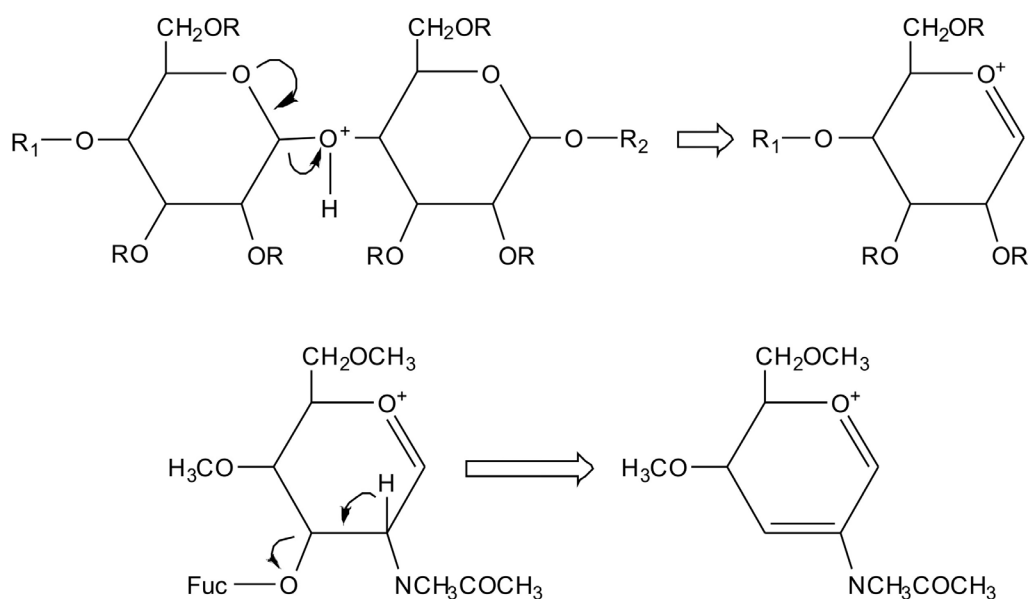
Figure 1.22 A type-cleavage yields an oxonium ion (top panel), β-elimination of the 3 position of the oxonium ion (bottom panel)

R_1 is the non-reducing end, R_2 is the reducing end. R stands for the 'remainder' of the chemical group.

1.5.5. Applications in glycomic analysis

The development of modern mass spectrometry, especially the revolution which occurred in the 1980s when MS could readily be used to analyse non-volatile samples, has facilitated glycomic analysis (Bern et al., 2013; Oates et al., 1985; North et al., 2010). Usually there are two goals in the glycomic analysis: structural determination and semi-quantitative analysis. These two goals can be achieved via the applications of MS.

1.5.5.1. Structural analysis

Not only can the structures of glycoproteins be determined by mass spectrometry (Dell and Morris, 2001; Morris et al., 1978), but also the glycans (Dell et al., 1983a; Babu et al., 2009; Jia et al., 2014). To determine structures of glycans in a sample, firstly the m/z values of these glycans are measured via MS. For MALDI-TOF instrumentation, the ions yielded are usually monocharged, therefore the m/z values obtained from the MS are actually the molecular weight of the glycans. Since the molecular weight is clear, possible monosaccharide compositions can be deduced. Moreover, the number of potential glycan

structures can be minimised either by using mass spectrometers with high mass accuracy or by the knowledge of the glycan biosynthesis pathway which has been well characterized (see sections 1.2.2 and 1.2.3) (Schwarz and Aebi, 2011; Varki, 2009; Weerapana and Imperiali, 2006; Dell et al., 2010).

N-glycan data from a glycomic paper is described below as an example of how to apply mass spectrometry in glycan structural analysis.

The N-glycan profile of ferret lung was obtained via MS analysis and glycans of interest were noted. For instance, one of the glycans, the m/z of which is 3252.4, has been determined to have the following composition Fuc1Hex4HexNAc6NeuAc2. The composition implies that the glycan may contain an Sda epitope (NeuAcα2,3(GalNAcβ1,4)Galβ1,4GlcNAc). Since MS analysis could not check the presence of the Sda, further investigation was carried out by breaking the glycan into fragment ions and then measuring the m/z of each fragment using MS/MS. The principle of MS/MS is that various glycans will yield different groups of fragment ions, and these fragments constitute a fingerprint of particular glycan structures. As shown in Figure 1.23, the most abundant fragment ion at m/z 2387.3 arose from cleavage of amino sugar glycosidic bonds. The signal at m/z 2183.2 corresponds to a loss of the tetrasaccharide constituent of the Sda capping group. Its concurrent ion is also observed at m/z 1092.5. This demonstrates the presence of Sda epitope (Jia et al., 2014). In addition, to make the result more accurate Jia et al. have also carried out GC-MS analysis on partially methylated alditol acetates derived from the permethylated glycans. The expression of the Sda epitope was confirmed by the presence of terminal GalNAc together with 3,4-linked Gal (Jia et al., 2014). Taken together, the results of MS, MS/MS and GC-MS have confirmed the presence of the Sda capping group in the sample.
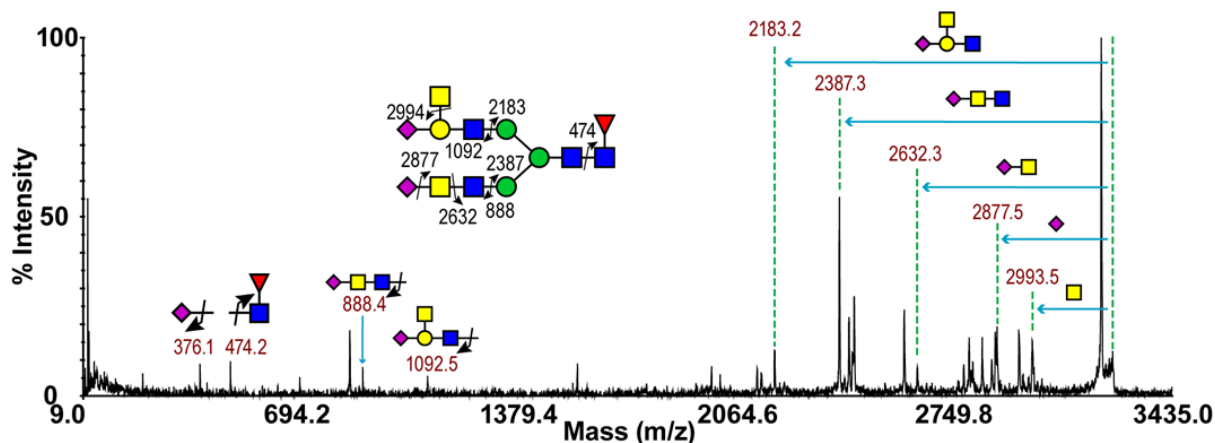
Figure 1.23 MALDI-TOF/TOF MS/MS spectrum of permethylated glycan at m/z 3252 which was derived from the ferret lung

Data were acquired in the form of $[M+Na]^+$ ions. The number indicated above the peak is the m/z value of the fragment ion (resulting ion) that has been detected by the mass spectrometry. Peaks were annotated with putative fragment ions according to the molecular weight (Jia et al., 2014). ■ GlcNAc, □ GalNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

1.5.5.2.   *Semi-quantitative analysis*

The reason why analysis is called semi-quantitative is that the MS spectrum shows relative abundance of each glycan but not the absolute amount. Although it is not fully quantitative, it has been reported that relative quantitation based on signal abundances of permethylated glycans investigated by MALDI-TOF MS is a reliable approach, especially when making a comparison of signals over a small mass range between glycans with similar structures within the same spectrum (Haslam et al., 2008; Wada et al., 2007).

In our laboratory, the most commonly used strategy in quantitative analysis is called label free. The label free approach together with mass spectrometry has been extensively employed in glycomic and proteomic studies (Neilson et al., 2011; Wada et al., 2007; Babu et al., 2009; Kailemia et al., 2014). MS spectra show not only m/z values of glycans but also their relative abundance. The relative abundance of different glycans is determined by label free quantification via comparing their molecular ion signals. It has been demonstrated that relative quantitation based on signal intensities of permethylated glycans analyzed by MALDI-TOF MS is a reliable method (Wada et al., 2007; Bateman et al., 2010). The permethylation leads to more consistent ionization due to the fact that chemical groups –OH, –COOH and –NHCOCH$_3$ on glycans have been replaced by –OCH$_3$, –COOCH$_3$ and – NCH$_3$COCH$_3$ respectively. In addition, for those sialic acids containing glycans which easily

66

lose sialic acids in MS in-source fragmentation, permethylation makes them more stable.

A glycomic study of neutrophils demonstrated the reliability of quantitative analysis of permethylated glycans: although the two normal samples are from two different countries, their glycan profiles are similar (Babu et al., 2009).

For semi-quantitative analysis of a mixture of glycan isomers, the characteristic fragment ions of each isomer can be observed during MS/MS analysis. In this case, the greater the percentage of an isomer in this mixture is, the higher the intensity of its characteristic ion will be, and this implies the relative abundance of the characteristic ions can be used to compare the amount of the isomers (North et al., 2010).

In this thesis I am employing mass spectrometry strategies in three projects which are comparative glycomic profiling of HeLa cells and normal human dermal fibroblast (NHDF), investigation of glycosylation in patients with muscular diseases, and glycomic profiling of human trophoblasts. Details of these projects are introduced successively in the following sections.

1.6. Ribosome-inactivating proteins (RIPs) and lectins from elderberry

1.6.1. RIPs

Ribosome-inactivating proteins (RIPs) are a group of proteins that function as protein synthesis inhibitors that enzymatically act at the ribosome (Stirpe, 2004; Stirpe, 2013). Since 1970 it has been reported that RIPs are less toxic to normal cells than cancer cells (Lin et al., 1970). Therefore RIPs have received a lot of attention as they may be potential antitumor therapeutic medicines. However, not all RIPs are suitable for being medicine candidates. Some RIPs, such as ricin and abrin, are not acceptable as they show powerful cytotoxicity towards mammalian cells, but some other RIPs are suitable as they possess strong protein synthesis inhibition activity *in vitro* but they are approximately 10^3-10^5 less toxic than ricin towards animal cells (Ferreras et al., 2011; Tejero et al., 2015). These less toxic RIPs include RIPs from *Sambucus*. Recently Van Damme and Shang at Ghent University found that lectins from *Sambucus nigra* (elderberry) are also less toxic to normal cells (Shang et al., 2015).

Although not all of the mechanism of cytotoxicity triggered by RIPs and lectins is completely clear, at least the well-studied RIPs can provide further impetus to study the physiological activities of RIPs and lectins on mammalian cells. More importantly, the way that RIPs and lectins from *Sambucus nigra* employed to show their cytotoxicity may be initiated by binding to cells, and this binding is probably achieved via their lectin chains.

The most intensively studied RIP by far is ricin. Hitherto, RIPs have been generally divided into two groups (Figure 1.24): type 1 RIPs, consisting of a single peptidic chain, the molecular weight of which is approximately 30 kDa, and type 2 RIPs which are composed of an enzymatically active A chain which is similar to type 1 RIPs, connected by a single disulphide bond to a B chain which has specific lectin properties (Stirpe, 2004; Lord et al., 1994; Stirpe, 2013; Tejero et al., 2015).

A RIP enzymatically cleaves one adenine base from the sugar-phosphate backbone of the rRNA via the N-glycosidase activity of the A chain, thus inhibiting protein synthesis (Endo and Tsurugi, 1987). Only specific adenines are targeted by the RIP. For example, if it is in rat liver, the adenine should be adenine 4324 which is located on a big loop of the 28S rRNA (Endo et al., 1987).
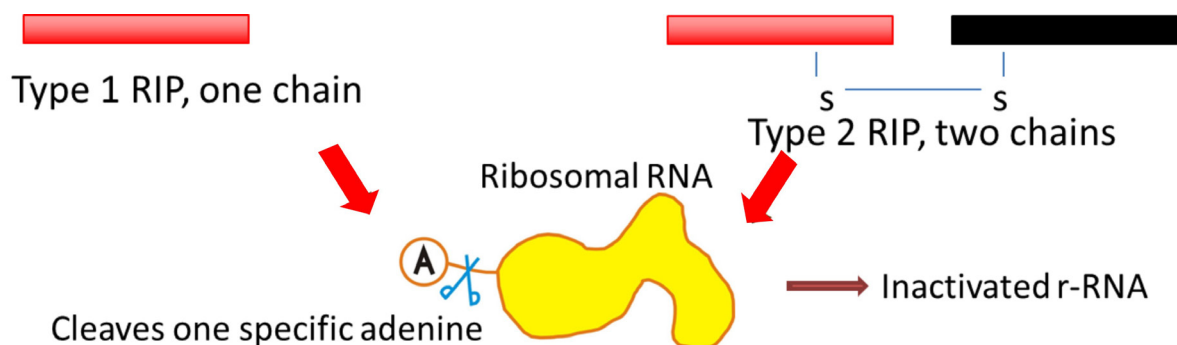
Figure 1.24 Schematic representation of the structures of two different types of RIP
The red bar stands for the one chain in type 1 RIP; the black bar is the chain B which has specific lectin properties. After one adenine (A) is removed, the r-RNA becomes inactivated.

In this project the RIPs that I focused on are SNA-I, SNA-V and SNALRP, which are classified as type 2 RIPs. Here SNA is short for *Sambucus nigra* agglutinin. For historical reasons the proteins are usually termed as SNA followed by Roman numbers attributed in

chronological order of their discovery (Shang and Van Damme, 2014). The full name of SNLRP is *Sambucus nigra* lectin related protein and this is due to the fact that it is closely related to type 2 RIP but it does not show agglutination activity (Van Damme et al., 1997a; Shang and Van Damme, 2014).

SNA-I is obtained from the storage parenchyma cells of elderberry bark and it is the most widely studied so far (Tejero et al., 2015). This protein can be extracted via affinity chromatography using immobilised fetuin. Analysis of SNA-I has demonstrated that this protein exists as a tetramer which is composed of four subunits, and each of the four subunits contains two disulphide bond linked chains. The molecular weight of the subunit is approximately 60kDa. Each subunit possesses eight putative N-glycosylation sites (Van Damme et al., 1997b; Shang and Van Damme, 2014). A main biological function of SNA-I is the RNA N-glycosidase activity. As its name implies, it can agglutinate animal and human erythrocytes. In addition, it also prevents protein synthesis and development of some insects. For instance, it has been reported that in the insect midgut SNA-I can trigger caspase 3-like protease-induced cell apoptosis (Van Damme et al., 1997b; Shahidi-Noghabi et al., 2011; Van Damme et al., 1998).

SNA-V, also termed Nigrin b, also exists in elderberry bark. This glycoprotein can be purified using affinity chromatography with immobilized GalNAc. It is a dimer which contains two [A-s-s-B] subunits, and each of them possesses six putative N-glycosylation sites. SNA-V also has RNA N-glycosidase activity, and thus can result in agglutination of rabbit and human erythrocytes (Van Damme et al., 1998; Tejero et al., 2015).

SNLRP is a predominant lectin in elderberry bark. It is a monomer which contains an enzymatically active A chain and a not fully active B chain, and thus it possesses RNA N-glycosidase activity but no agglutination activity. It has five potential N-glycosylation sites. This is probably due to the fact that there are several amino acid substitutions in the B chain sugar binding sites, and these substitutions have affected the binding affinity (Van Damme et al., 2001).

The parameters of the above three type 2 RIPs are summarized in the following table (Shang and Van Damme, 2014).

Table 1.4 Overview of type 2 RIPs from *Sambucus nigra*

Name	Source	Structure	Molecular weight	Protein synthesis inhibition	Agglutination activity
SNA-I	Bark	Tetramer $[A\text{-}s\text{-}s\text{-}B]_4$	240kDa (4×60kDa)	+	+
SNA-V	Bark	Dimer $[A\text{-}s\text{-}s\text{-}B]_2$	120kDa (2×60kDa)	+	+
SNLRP	Bark	Monomer $[A\text{-}s\text{-}s\text{-}B]$	62kDa	+	-

1.6.2. Lectins

Lectins are specific carbohydrate-binding proteins. They are ubiquitous, being identified in microorganisms, plants and animals (Berg et al., 2002b). *Sambucus nigra* lectins can be classified as R-type which is a superfamily of glycan-binding proteins that contain a carbohydrate-recognition domain which is structurally similar to that in ricin (Varki, 2009). The difference between the lectin from *Sambucus nigra* and the RIP is the former only has a lectin chain.

In this project the lectins that I am interested in are SNA-II and SNA-IV. An overview of these two lectins is shown in Table 1.5 (Shang and Van Damme, 2014). The nomenclature of these lectins is the same as the one used in the RIPs.

SNA-II is one of the most abundant lectins in elderberry bark. It can be purified using affinity chromatography with immobilized asialoglycophorin. This SNA actually shares a precursor with SNA-V. The precursor is processed in two different ways, as a result of which two different products, SNA-II and V, are yielded. SNA-II is a homodimer containing two B chains which are the same as the B chain in SNA-V without the initial eight amino acids. Each of the B chain has four putative N-glycosylation sites (Shahidi-Noghabi et al., 2011). Due to the lack of an A chain, SNA-II is not able to show RNA N-glycoside activity and thus it cannot inhibit protein synthesis (Van Damme et al., 1998; Maveyraud et al., 2009).

Unlike SNA-II, SNA-IV is a predominant lectin in the elderberry fruit. It can be purified using affinity chromatography with immobilized GalNAc. Like SNA-II, it shows agglutination activity but it does not exhibit RNA N-glycosidase activity and cannot inhibit protein synthesis. There are three putative N-glycosylation sites in each B chain.

Table 1.5 Overview of lectins from *Sambucus nigra*

Name	Source	Structure	Molecular weight	Protein synthesis inhibition	Agglutination activity
SNA-II	Bark	Dimer [B]$_2$	60kDa (2×30kDa)	-	+
SNA-IV	Fruit	Dimer [B]$_2$	64kDa (2×32kDa)	-	+

Lectins of the ricin B chain form are composed of two β-trefoil domains which are the binding domains and contain 3 subdomains, α, β and γ. The amino acid residues that form the glycan-binding site in ricin have been determined via cocrystalization of ricin with several glycan structures (Montfort et al., 1987). It has been reported that the α-subdomain of the first β-trefoil domain and the γ-subdomain of the second β-trefoil domain are involved in the glycan binding (Rutenber and Robertus, 1991).

The result of sequence alignment of the ricin B chain and the RIP B chains and lectins from *Sambucus nigra* shows that in both subdomain-Iα and subdomain-IIγ three of five residues which form the carbohydrate binding site are conserved between ricin and all SNAs (Table 1.6) (Shang and Van Damme, 2014). This suggests that in respect of glycan binding the B chains of SNAs show similarity to that of ricin.

Table 1.6 Comparative analysis of the residues forming the carbohydrate binding sites of the five SNA lectin chain and ricin lectin chain

RIP B chain/lectin	Subdomain-Iα	Subdomain-IIγ
SNA-I	Asp26, Gln39, Arg41, Asn48, Gln49	Asp231, Ile243, Tyr245, Asn252, Gln253
SNA-II	Asp16, Gln29, Trp31, Asn38, Gln39	Asp227, Ile239, Phe241, Asn248, Gln249
SNA-IV	Asp18, Gln31, Trp33, Asn40, Gln41	Asp229, Ile241, Phe243, Asn250, Gln251
SNA-V	Asp24, Gln37, Trp39, Asn46, Gln47	Asp235, Ile247, Phe249, Asn256, Gln257
SNLRP	Asp23, Gln36, Leu38, Ser45, Gln46	Glu230, Ile242, Tyr244, Asn251, Gln252
Ricin	Asp15, Gln27, Trp29, Asn39, Gln40	Asp227, Ile238, Tyr241, Asn248, Gln249
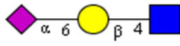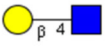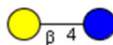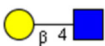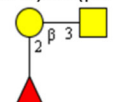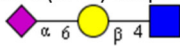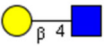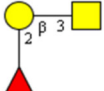
1.6.3. Glycan binding

Van Damme and Shang have carried out glycan array analyses of these five proteins. The result showed that SNA-I exhibited strong binding affinity to glycans containing α2,6 linked terminal sialic acid, and all strongly interacting glycans contain at least one this sialic acid. SNA-I exhibited the strongest binding to the glycan array among all SNAs, and this is consistent with its very strong interaction with cells and glycoproteins. SNA-II, IV and V all

interacted with glycans possessing Gal or GalNAc. Unlike SNA-II and V, SNA-IV also binds to sialylated Gal residues. SNLRP binds to glycans containing GlcNAc (Shang and Van Damme, 2014). The top three glycan motifs that reacted with these five SNAs have been summarized in the following table (Shang and Van Damme, 2014) to show the SNA binding preference.

Table 1.7 Top three glycan motifs that reacted with *Sambucus nigra* type 2 RIPs and lectins
■ GlcNAc, □ GalNAc, ● Glc, ○ Gal, ◆ NeuAc, ▲ Fuc.

Protein		1. Glycan motifs	2. Glycan motifs	3. Glycan motifs
Type 2 RIP	SNA-I	Neu5Ac(α2-6)-Galβ1-4GlcNAc	Gal (β1-4)GlcNAcβ	Gal(β1-4)-Glcβ
	SNA-V	Gal(β1-4)-Glcβ	Fuc(α1-2)Gal(β1-4)GlcNAcβ	GalNAc(β1-4)GlcNAcβ
	SNLRP	GlcNAc(β1-4)GlcNAcβ	Gal (β1-4)GlcNAcβ	
Lectin	SNA-II	Fuc(α1-2)Gal(β1-4)GalNAcβ	Gal(β1-4)-Glcβ	GalNAc(β1-4)GlcNAcβ
	SNA-IV	Neu5Ac(α2-6)-Galβ1-4GlcNAc Neu5Ac(α2-3)-Galβ1-4GlcNAc	Gal(β1-4)GlcNAcβ	Fuc(α1-2)Gal(β1-4)GalNAcβ

1.7.   Muscular diseases caused by gene mutations

Gene mutations are changes occurring in the genetic sequence, and they can have different consequences. One of the consequences is that gene mutations can result in various diseases, such as muscular diseases. According to the clinical features of the muscular diseases mentioned in the papers (Guergueltcheva et al., 2012; Senderek et al., 2011; Palace, 2012; Palace et al., 2007; Downham et al., 2008; van der Ploeg and Reuser, 2008), this group of mutations can be classified as 'loss of function' type, the effect of which is that the gene products are complete or partial loss of function. The phenotypes connected with these

mutations are usually recessive. Many of the genes, such as *GFPT1*, *DPAGT1*, *ALG2* and *ALG14*, play important roles in the synthesis of dolichol-linked oligosaccharide, and the mutations in these genes result in congenital myasthenic syndromes (Freeze et al., 2015; Senderek et al., 2011; Cossins et al., 2013; Belaya et al., 2012). Although the gene mutation sites and most of carbohydrate metabolism have already been detected, our knowledge and understanding of these diseases is limited, especially in the glycobiology field. For instance, it is unknown why, although the phenotypes of CDG and myasthenic patients are different, they could be caused by mutations in a same gene, such as *GFPT1* (Freeze et al., 2015).

1.7.1. Congenital myasthenic syndromes (CMS) caused by *GFPT1* mutations

In the past few years mutations in genes encoding protein glycosylation enzymes, or enzymes synthesising the building blocks for protein glycosylation, have been identified to cause a neuromuscular transmission defect named congenital myasthenic syndromes (CMS) which are autosomal recessive diseases that characterized by a limb-girdle pattern of muscle weakness and usually accompanied by the presence of tubular aggregates in muscle biopsies (Senderek et al., 2011; Guergueltcheva et al., 2012). The first of these genes to be correlated with CMS was *GFPT1* (glutamine-fructose-6-phosphate transaminase 1), which has been recently identified by a group of researchers including Lochmuller and Mueller at Newcastle University using classical positional cloning (Senderek et al., 2011; Guergueltcheva et al., 2012). Subsequently, mutations in two genes (*ALG2* and *ALG14*), encoding enzymes of the protein N-glycosylation pathway were also found to cause CMS (Cossins et al., 2013). CMS caused by *GFPT1* mutations is a type of neuromuscular junction disorder in which the synapses that form between motorneurons and skeletal muscle fibres that transmit the impulse resulting in muscle contraction have impaired function (Senderek et al., 2011; Sanes and Lichtman, 1999; Martin, 2002; Guergueltcheva et al., 2012). Synapses are the fundamental units of the human nervous system (Martin, 2002; Senderek et al., 2011), which implies that synapses could play an important role in the function of the neuromuscular junction. Research from 40 years ago suggested that some glycoproteins, such as acetylcholine receptors and acetylcholinesterase, are concentrated at neuromuscular synapses (Sanes and Cheney, 1982; Conti-Tronconi and Raftery, 1982; Massoulie and Bon, 1982), and this implies that glycoproteins are essential for synaptic function or development. In addition,

many vital neuromuscular proteins are glycosylated, such as agrin and dystroglycan (Senderek et al., 2011; Martin, 2002). Thus attention should be paid to the glycosylation in the neuromuscular junction. However, most neurobiologists usually ignored the glycosylation when they were thinking about synapses, thus research on synaptic function has been carried out using approaches in which glycans were not considered (Martin, 2002), as a result of which they would narrowly miss an opportunity to discover something of interest in the synapse. Nowadays, some neurobiologists have realized this problem.
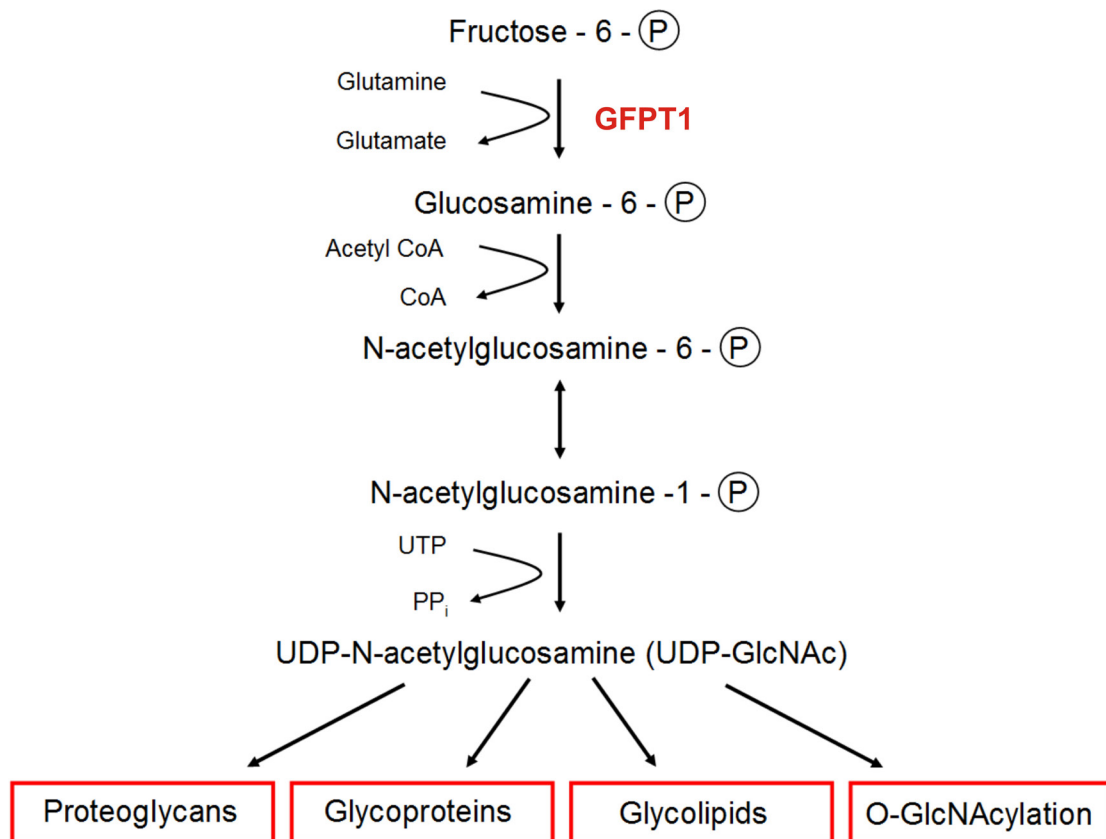
Figure 1.25 A simplified hexosamine biosynthesis pathway

Usually less than 5 % of the fructose-6-phosphate isomerized takes part in hexosamine pathway. The first and rate-determining step is coupling of an amino group from glutamine (Gln) to fructose-6-phosphate. This reaction is catalysed by GFPT1 or GFAT1. The product Glucosamine-6-phosphate then reacts with acetyl-CoA to form N-acetylglucosamine-6-phosphate. After that, an isomerase change N-acetylglucosamine-6-phosphate into N-acetylglucoseamine-1-phosphate. Finally, the latter reacts with UTP, producing UDP-N-acetylglucosamine (UDP-GlcNAc) which acts as a precursor for amino sugar used for synthesis of proteoglycans, glycoproteins, glycolipids and O-GlcNAc (Senderek et al., 2011).

GFPT1 (glutamine-fructose-6-phosphate transaminase 1), also named GFAT1 (glutamine-fructose-6-phosphate aminotransferase 1), is the first enzyme of the hexosamine biosynthesis pathway which has been shown in Figure 1.25 (UniProtKB, 2012; Senderek et al., 2011). It

transfers an amino group from glutamine to fructose-6-phosphate to yield glucosamine-6-phosphate, providing the precursor for uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) synthesis. UDP-GlcNAc is an essential substrate for all mammalian glycosylation biosynthesis pathways and N-glycan branching is especially sensitive to alterations in the concentration of this sugar nucleotide (Senderek et al., 2011; UniProtKB, 2012; Freeze et al., 2015).

In GFPT1, there are two sugar isomerase (SIS) domains and one glutamine amidotransferase domain (UniProtKB, 2012; Senderek et al., 2011). The active site of this protein is still unclear though there is a prediction about the active site (UniProtKB, 2012). GFPT1 has two isoforms: isoform 1 and isoform 2 which are composed of 699 amino acids and 681 amino acids, respectively. The sequence difference between isoform 1 and isoform 2 is that peptide 229-246 is missing in the latter (UniProtKB, 2012).

In 2011, 16 CMS patients caused by *GFPT1* mutations were reported (Senderek et al., 2011). In the next year, more CMS patients were discovered and the clinical features of all these patients were further documented (Guergueltcheva et al., 2012). In 2013, 12 novel mutations in CMS patients were identified (Selcen et al., 2013). The details of the *GFPT1* CMS patients that are associated with this project are listed in the following table. The mutation sites in the other four patients are not clear.

Table 1.8 The country of origin, gene mutation sites and the amino acid changes caused by the mutations

CMS	Country of origin	Nucleotide change	Amino acid changes
GFPT1 patient 1	Germany	595G>T	Val199Phe
GFPT1 patient 2	Spain	1475T>C	Met492Thr
GFPT1 patient 3	Spain	1475T>C	Met492Thr

Having discovered that some of their CMS patients had mutations in *GFPT1*, Lochmuller et al. had found that these patients have a recognizable pattern of weakness of shoulder and pelvic girdle muscles, and defects in ocular or facial muscles showing a favourable treatment response to acetylcholinesterase (AChE) inhibitors. They discovered that the enzymatic activity of control GFPT1 did not change too much compared that of the mutants (Figure 1.26A). The comparison of GFPT1 amounts in patients and the control showed that the amount of GFPT1 in the two patients reduced to 51% and 22% of the GFPT1 amount in the control (Figure 1.26B) (Senderek et al., 2011). In addition, they also set up an animal model using zebrafish. The result from the animal model showed that low expression level of

GFPT1 resulted in delayed neuromuscular junction development and abnormal muscle morphology (Senderek et al., 2011), GFPT1 was found to be essential for events in neuromuscular transmission. *GFPT1* knockout mice are not available yet.

A
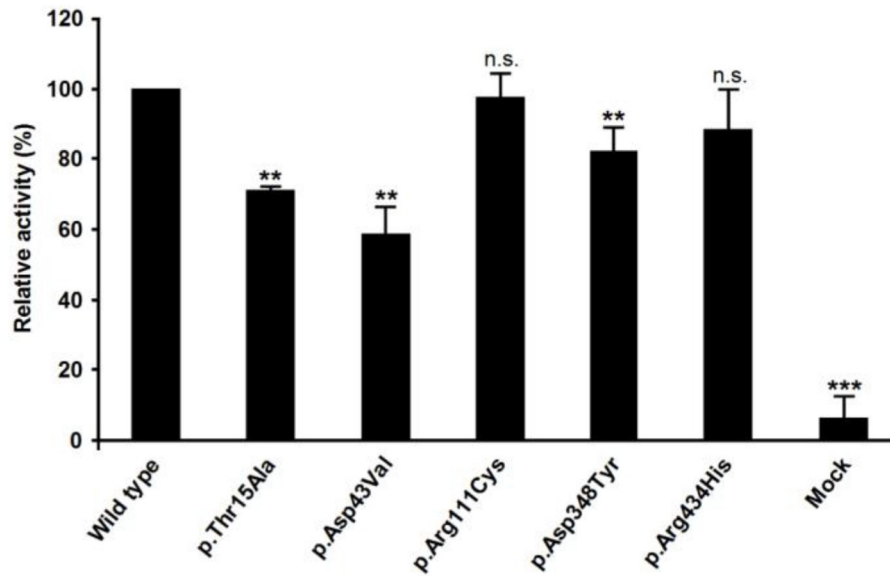
B
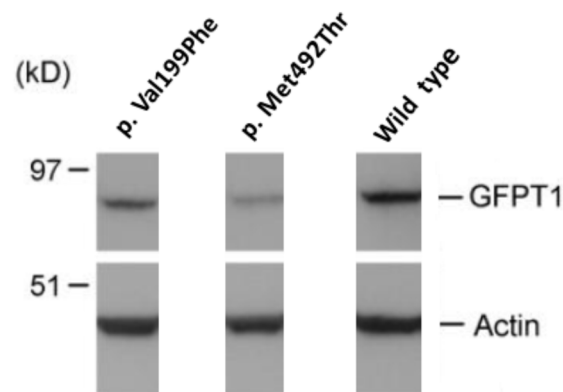
Figure 1.26 Analysis of the enzymatic activity of *GFPT1* mutants (A), Western blot analysis of GFPT1 expression in myoblasts (B)

In A, HEK293 cells were transfected with either wild type or mutant *GFPT1* constructs. Enzyme activity of GFPT1 was measured in cell lysates with the glutamate dehydrogenase method 48 h after transfection. The enzymatic activity of every mutant was normalized to GFPT1 protein amounts determined by western blot analysis of cell lysates used for activity measurements. Triplicate experiments were carried. Error bars indicate ± SD; significant differences from wild type $**P < 0.01$; $***P < 0.001$. n.s., not significant (Senderek et al., 2011). In B, cell lysates of cultured myoblasts and differentiated myotubes of patients 3 and 5 (shown in Table 1) and one healthy control donor were immunoblotted with an anti-GFPT1 antibody (upper panel). Actin was used as a control. GFPT1 band intensities were normalized to actin bands in the same lane (Senderek et al., 2011).

Since they have not carried out any glycomic investigation on *GFPT1* patients, and GFPT1 plays an essential role in synthesizing UDP-GlcNAc which is involved in the initiation of N-glycan antennae (Figure 1.27), we hypothesized that mutations in *GFPT1* may cause N-glycan branching variations and thus have an influence on protein glycosylation. This hypothesis can be tested via a systematic glycomic study on *GFPT1* patients. Additionally, samples from *DOK7* CMS patient and other muscular disease patients were also investigated. This was to check if there was a glycosylation variation and whether this variation was unique in the *GFPT1* CMS patients.
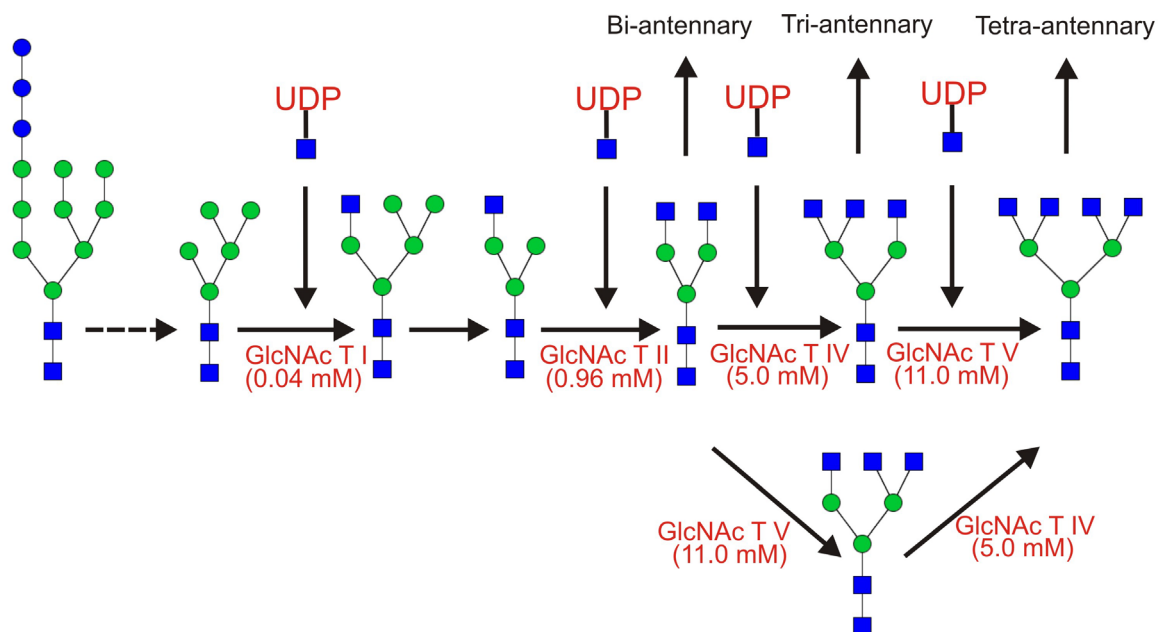
Figure 1.27 UDP-GlcNAc is involved in the initiation of N-glycan antennae
■ GlcNAc, ● Man, the numbers in the brackets under the enzyme are the K_m values (Lau et al., 2007).

1.7.2. Muscular diseases caused by other gene mutations

Mutations occurring in other genes which are not encoding protein glycosylation enzymes, or enzymes synthesising the building blocks for protein glycosylation can also result in muscular diseases.

1.7.2.1. CMS caused by DOK 7 mutations

DOK7 encoded by *DOK7* is a postsynaptic protein involved in the acetylcholine receptor (AChR) clustering pathway (Palace, 2012). It functions as an activator of muscle, skeletal receptor tyrosine kinase (MuSK) via dimerization (Bergamin et al., 2010). It is likely that it plays an important role in neuromuscular synaptogenesis.

In 2006, *DOK7* mutation was firstly reported as a cause of CMS (Beeson et al., 2006). One year later, this type of mutation has been further shown to underlie a recessive CMS associated with small simplified neuromuscular junctions but normal acetylcholine receptor and acetylcholinesterase function (Palace et al., 2007). Onset of DOK7 CMS tends to occur at the ages of two to three years (Palace, 2012). The characteristic of the onset is that patients show difficulty in walking development after normal motor milestones (Palace et al., 2007). Proximal muscles were severely affected, resulting in weakness in limb-girdle. It is common that the symptoms become worse in adulthood, particularly impacting respiratory function. Treatment with ephedrine or oral salbutamol can result in a slow, steady, and often dramatic improvement over months (Palace, 2012). These characteristics of *DOK7* CMS are different from limb-girdle myasthenia associated with tubular aggregates, in which *GFPT1* mutations were detected and patients showed positive response to anticholinesterase treatments (Palace et al., 2007; Senderek et al., 2011).

It has been reported that *DOK7* mutation is the third most common cause of CMS in the UK (Palace, 2012). Our knowledge and understanding of this disease is limited, thus investigation will help clinical diagnosis and potential treatment.

1.7.2.2. *Muscular disease caused by MTND5 mutations*

MTND5 is short for mitochondrial *NADH dehydrogenase 5* gene which encodes the subunit 5 of the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I). Like the position of GFPT1 in hexosamine pathway, Complex I is the first enzyme in the mitochondrial electron transport chain (Liolitsa et al., 2003). In eukaryotes, Complex I is situated in the inner membrane of the mitochondria. It functions as a catalyst to transfer electrons from NADH to coenzyme Q10.

It has been reported that mutations in *MTND5* can result in neurodegenerative disorders, such as Leigh syndrome and Parkinson's disease (Parker and Parks, 2005; Bannwarth et al., 2013;

Petruzzella et al., 2003). However, researchers also find that *MTND5* mutations can lead to muscular diseases, such as myopathies (Alston et al., 2010; Downham et al., 2008).

Recently a young girl with defect in Complex I in muscle caused by a mutation, m.12425delA, within the *MTND5* gene was reported. This single deletion was not detected in her mother but identified in some tissues from her proband, which implies that this is a *de novo* mutation. Her cardinal symptoms are renal failure and myopathy (Alston et al., 2010).

1.7.2.3. *Limb girdle muscular dystrophy type 2A (LGMD2A) caused by CAPN3 mutations*

LGMD2A is considered to be the most common form of recessive LGMD. It is an autosomal recessive disorder caused by mutated *CAPN3* gene which encodes deficient calpain-3 protein (calcium-activated neutral protease 3). Calpain-3 protein is a cysteine protease belonging to the intracellular calpain family (Rocha and Hoffman, 2010; Richard et al., 1995). It is a muscle-specific enzyme. Although the role of this enzyme plays in normal skeletal muscle physiology and in LGMD2A pathophysiology is not clear (Richard et al., 1995; Fanin and Angelini, 2015), a study suggests that it is important for a physiological process in muscle termed "sarcomere remodelling" (Kramerova et al., 2005).

Currently the way to diagnose LGMD2A is detecting calpain-3 protein deficiency in muscle and then identifying the causative mutations. Other biochemical tests are not sensitive (Fanin and Angelini, 2015; Rocha and Hoffman, 2010).

1.7.2.4. *Pompe disease caused by GAA mutations*

Pompe disease is a disorder of glycogen accumulation. It is caused by mutations in the gene *GAA* which encodes acid α1,4 glucosidase, a lysosomal enzyme. This enzyme plays an essential role in the decomposition of glycogen in lysosomes. Its deficiency mainly results in lysosomal glycogen over storage, usually in cardiac and skeletal muscles, due to the incapacity of breaking down glycogen into Glc (Matalon et al., 2006; Turaca et al., 2015).

In the process of translation, lysosomal enzymes enter the ER where glycosylation occurs co-translationally (Hermans et al., 1993). Research shows that a suitable protein conformation is indispensable for recognition and glycosylation of a site (Bause and Legler, 1981). After the three terminal Glc residues are removed from the 14-sugar (2 GlcNAc, 9 Man and 3 Glc) precursor in the ER, a Man-6-phosphate recognition marker is acquired as a lysosomal targeting sign by most of the lysosomal enzymes. This is achieved by transferring GlcNAc-1-phosphate from UDP-GlcNAc to a specific mannose residue, and subsequently the phosphate is released by a phosphodiesterase. Combining with the mannose 6-phosphate receptor this process is vital for lysosomal targeting. The enzyme is then shipped to the endosomes in which a pH decrease results in the dissociation of the ligand and the receptor. The enzyme continues its way to the lysosome and the receptor goes back to the Golgi apparatus (Hermans et al., 1993). It is known that the acid α1,4 glucosidase will follow this pathway and undergo modification with Man-6 phosphate substituted N-glycans. It possesses 7 N-glycosylation sites. Importantly, removal of the second glycosylation site at Asn-233 affects significantly the formation of mature enzyme (Hermans et al., 1993). However, it is not clear whether the mutation occurring in this gene had altered the cell glycome.

1.8. The protection of the embryo/foetus from the maternal immune system

1.8.1. Background

In 1953 Sir Peter Medawar raised a question 'how can the mother nourish a foetus which is similar to a "foreign organ" (paternal contribution to its half genome) within itself for several months but not reject it?' He gave three possible explanations: the antigenicity of the foetus is immature; maternal immunological indolence or inertness, and anatomical separation from the mother (Medawar, 1953). However, Medawar's hypotheses could not be supported by subsequent investigations. Immunity can be induced in skin transplantation tests when injecting foetal tissue (Billingham et al., 1956), arguing against the first possible explanation; the mother shows response to pathogens in pregnancy (Head and Billingham, 1986), which means the second one has been denied; the extravillous cytotrophoblast (evCTB) can invade the maternal decidua and myometrium and remodel the uterine arteries for the foetus

(Handschuh et al., 2007; Clark and Schust, 2013; Benirschke, 1994), and this is in conflict with the last one.

1.8.2. The human fetoembryonic defense system hypothesis (hu-FEDS)

To address the question, in 1996 Gary F. Clark et al. raised a hypothesis termed the human fetoembryonic defense system hypothesis (hu-FEDS) (Clark et al., 1996), and later it was renamed as eutherian fetoembryonic defence system hypothesis (eu-FEDS) to apply more broadly. It is a hypothetical model depicting a way via which the human immune system is able to recognize foreign substance as "own species" as has been observed with maternal immune tolerance in pregnancy. The fundamental idea of this hypothesis is that glycoproteins, which exist in the reproductive system and are exposed on gametes can either inhibit immune responses or prevent rejection of the foetus (Clark et al., 2001). Biomolecules in human seminal plasma and the pregnant uterus had been shown to suppress immune responses *in vitro* (Bolton et al., 1987; Kelly and Critchley, 1997), but the idea that human gametes could have an influence on immune responses was novel. This model has been intensively tested since it was conceived, and an increasing amount of supportive evidence for this model has been collected (Kui Wong et al., 2003; Pang et al., 2007). The following section talks about the gametes, from which the foetus initially develops.

1.8.3. Human gametes and maternal immune responses

In humans the gametes are sperm and eggs. Most of the information available suggests that sperm and eggs do not express detectable amounts of human leukocyte antigens (HLA) on the surface (Hutter and Dohr, 1998; Clark, 2014). If they did, the mother would reject the paternal sperm as a foreign organ transplant. The genes encoding HLA antigens are located on chromosome 6p21. HLA class I genes have been divided into two types, class Ia and class Ib which contains HLA-A, -B, -C, and HLA-E, -F, and -G, respectively. HLA class II (HLA-D) genes are not translated in human trophoblast cells (Murphy and Tomasi, 1998; Hunt et al., 2005; Apps et al., 2009). The cytotoxic T lymphocyte (CTL) is responsible for this HLA rejection as it kills cells with foreign HLA. Since sperm does not express detectable amounts

of HLA, it is 'invisible' to CTL. However, in the mother there are other immune cells, such as natural killer (NK) cells, which can eliminate cells without detectable HLA class I molecules (Karre, 2002; Jaeger and Vivier, 2012). Therefore, it is likely that the sperm will be killed by NK cells. However, studies have shown that human sperm express bisecting type N-glycans which could inhibit NK cell cytotoxicity (Pang et al., 2007; Patankar et al., 1997). Indeed, multiple pieces of evidence have shown that bisecting type N-glycan can suppress NK cell cytotoxicity. For instance, human K562 erythroleukaemic cells are easily killed by natural killer cells. However, stable transfection of K562 cells with the gene that encodes GlcNAc-transferase III which is responsible for transferring a GlcNAc to the bisecting site could produce NK cell resistant transformants (Patankar et al., 1997; Yoshimura et al., 1996). It is therefore likely that the sperm can evade these two major immune cells via glycan-mediated processes.

Recent studies also suggest that mechanisms to avoid the human immune system used by gametes, tumour cells and pathogens may have something in common. To some extent, they display molecular mimicry which is the expression of surface structures similar to those found in the host (Kui Wong et al., 2003; Pang et al., 2007; Clark, 2014).

1.8.4. Human foetus and maternal immune responses

A zygote forms as a result of fertilization, and then it becomes an embryo which develops into a foetus. The foetus floats in the amniotic fluid in the uterus, and it communicates with the mother via an interface termed placenta. Foetal development *in utero* requires a functional placenta that mediates the transport of nutrients and gases from the maternal blood which are essential for maintaining foetal viability. In addition to this, its development also requires the protection of the embryo/foetus from the maternal immune system. The blood circulations of the foetus and the mother are separate (Hunt et al., 2005). The foetus is like a foreign transplant organ for the latter due to the fact that it expresses paternal HLA molecules (Guleria and Sayegh, 2007). The pregnant uterus yields two major glycoprotein molecules, CA125 and glycodelin-A (GdA), which can suppress the potential maternal immune response against the foetus (Lee et al., 2009; Kui Wong et al., 2003; Clark and Schust, 2013), so the foetus is safe. Clinically, the major part of the placenta is composed of foetal elements; only a very small amount of maternal decidua will be shed at delivery (Benirschke, 1998). This

suggests that for the mother the placenta is foreign. Anatomically, a placenta is composed of three principal parts: trophoblasts, connective tissue with chorionic membrane and blood vessels, and the amnion (Benirschke, 1998). Among these 3 parts, the trophoblast is in very intimate contact with the maternal blood in which the immune cells can circulate (Benirschke, 1998; Juch et al., 2012), which suggests that it is the placenta trophoblast that would be the target for the maternal immune cells. The attack would normally be expected to result in immune responses, and this could trigger severe pregnancy complications. However, the fact is that most pregnant women do not suffer these pregnancy complications, which implies that the immune response has been suppressed. The underlying mechanism of how the immune response is suppressed at the placenta trophoblasts remains unknown. It is therefore necessary to pay attention to the trophoblasts.

Placenta trophoblasts (TB) are originally derived from the trophectoderm, the outer layer of a blastocyst (Georgiades et al., 2002; Paria et al., 2002). They are considered to be trophoblastic stem cells as they can differentiate into two different types of cells: the villous cytotrophoblast (CTB) and the extravillous cytotrophoblast (evCTB). The former fuses with each other, forming syncytiotrophoblasts (STB) which are multinucleated. CTB and STB stay on the foetal side (Figure 1.28), and they cover the chorionic villi and are responsible for the exchange of gases and nutrients between the mother and the foetus via the blood (Tarrade et al., 2001; Georgiades et al., 2002; Handschuh et al., 2007; Ji et al., 2013). The latter, evCTB, extravasates and invades the maternal decidua and myometrium where they act with decidual NK cells to remodel uterine arteries into flaccid conduits. Thus evCTB cells are spread out on the maternal side (Figure 1.28). Defects in the invasion can result in pregnancy complications, including late miscarriage and preeclampsia (Handschuh et al., 2007; Clark and Schust, 2013; Benirschke, 1994; Robson et al., 2012; Parham, 2004).
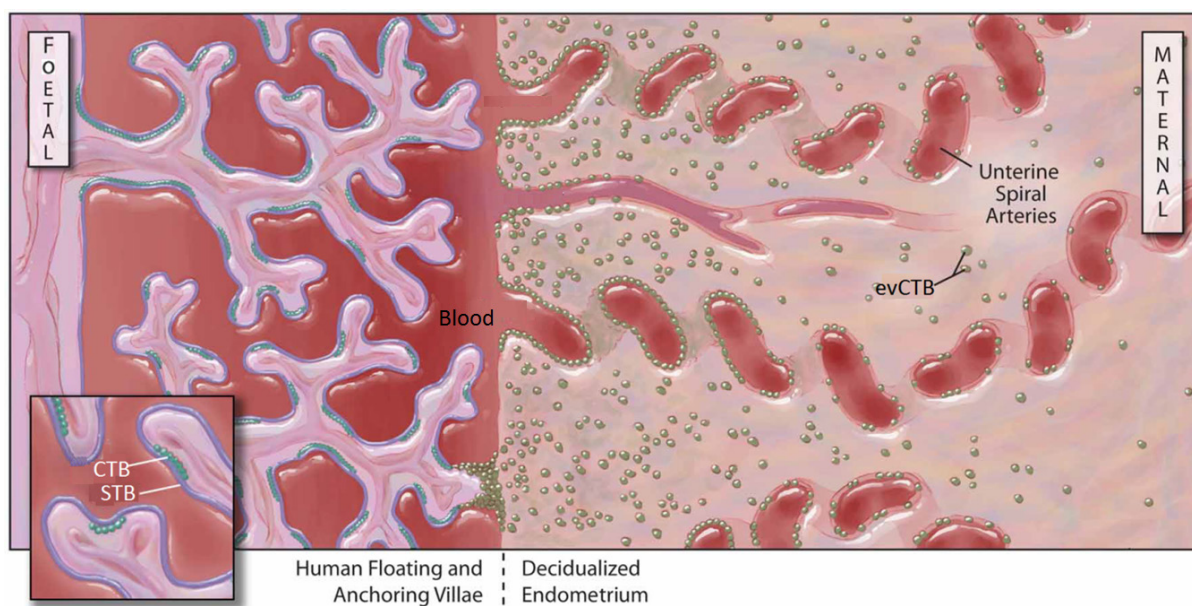
Figure 1.28 Human placental plate structure after 12 weeks of gestation: the placenta has a foetal and a maternal side

The foetal part is composed of branched villous structures which are soaked in maternal blood for nutrition and gas exchange. The maternal decidua is full of stromal and immune cells and is traversed by spiral arteries which pump blood into the intervillous space. The placental villi are covered by an inner layer of mononucleated villous cytotrophoblasts (CTB) and an outer layer of fused syncytiotrophoblasts (STB). The extravillous cytotrophoblasts (evCTB) are extravasated (green spots) and then spread out in the maternal part (Clark and Schust, 2013).

1.9. Aims of this thesis

The main aim is to investigate biomedically important glycosylation using mass spectrometric approaches.

There are three specific aims:

- Glycomic analysis of HeLa cells and NHDF for developing potential antitumor therapeutic medicines (Chapter 3).
- Glycomic investigation of CMS caused by *GFPT1* mutations and other muscular diseases caused by other gene mutations (Chapter 4).
- To further understand the role of glycans in human fetoembryonic defence system by structurally characterizing trophoblast N-glycans using mass spectrometry (Chapter 5).

Chapter 2

Materials and methods

2. Materials and methods

2.1. Materials

2.1.1. General chemicals and reagents

The ultra-pure water (Neptune Purite water purification system, Purite Ltd, Oxfordshire, UK), nitrogen and argon (BOC, Guildford, UK), sodium chloride (NaCl) (Rose Chemicals Ltd, London, UK) and 3,4-diaminobenzophenone (DABP) (Acros Organics, Geel, Belgium). Methanol, ammonia, dimethylsulphoxide (DMSO), propan-1-ol, acetic acid, acetonitrile, chloroform, butanol, sodium hydroxide and trifluoroacetic acid (TFA) are from Romil (Cambridge, UK). Idoacetic acid (IAA), tris(hydroxymethyl)aminomethane (Tris), α-cyano-4-hydroxycinnamic acid, sodium cholate, hexane, potassium borohydride (KBH_4), 2,5-dihydroxybenzoic acid (DHB) and sodium borodeuteride ($NaBD_4$) are from Sigma-Aldrich (Poole, UK). Ammonium bicarbonate (AMBIC, NH_4HCO_3), 3-(N-Morpholino)-propanesulfonic acid (MOPS), manganese chloride 4-hydrate, UDP-Gal, sodium acetate (CH_3COONa), ethylenediaminetetraacetic acid (EDTA), potassium hydroxide (KOH), formic acid and Dowex 50WX8 beads are from Fluka (Polle, UK). 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate (CHAPS) and dithiothreitol (DTT) are from Roche (East Sussex, UK). Acetic anhydride and methyl iodide are from Alfa Aesar (Lancaster, UK).

2.1.2. Standards and enzymes

Calibration standards for MS and MS/MS are 4700 calibration standard kit (AB Sciex, Warrington, UK), glycan standards (Dextra, Reading, UK).

Porcine pancreas trypsin (Sigma-Aldrich, Poole, UK), PNGase F (*Flavobacterium meningosepticum*, recombinant from *E. coli*, Roche, East Sussex, UK), rEGCase II (*E. coli* encoding the gene of this enzyme yielded by *Rhodococcus sp.*, Takara, Saint-Germain-en-Laye, France), β1,4-galactosyltransferase (Merk, Darmstadt, Germany), endo-β-galactosidase

(*Escherichia freundii*, AMSBIO, Oxford, UK), sialidase S (expressed in *E. coli*, recombinant from *Streptococcus pneumonia*, Prozyme, Cambridge, UK) and sialidase A (expressed in *E. coli*, recombinant from *Arthrobacter ureafaciens*, Prozyme, Cambridge, UK).

2.1.3. Biological samples

All samples were kindly provided by my collaborators, the samples were then subjected to glycomic analysis in our laboratory.

2.1.3.1. NHDF (normal human dermal fibroblasts) and HeLa cells

Two batches of NHDF and HeLa cells were cultured *in vitro* by Professor Els Van Damme and Miss Chenjing Shang at Ghent University, Belgium.

2.1.3.2. Myoblasts and myotubes

Three batches of samples were provided by Professor Hanns Lochmuller and Dr. Juliane Mueller at Newcastle University. The first two batches are myoblasts; the last batch is the myotube.

In the first batch there were five myoblast cell lines which were derived from the muscle biopsies obtained from five separate patients and then immortalised *in vitro*. These patients are one *DOK7* patient, one *MTND5* patient, *GFPT1* patients 1, 2 and 3. Each cell line except *GFPT1* patient 3 has three pellets which were harvested from myoblasts cultured in the skeletal muscle cell growth medium containing 5%, 10% and 15% foetal calf serum (FCS). The *GFPT1* patient 3 cells grew extremely slowly and they only grew in the medium with 15% FCS.

In the second batch, there were nine myoblast cell lines. In addition to the five mentioned in the first batch, the other four cell lines were derived from the muscle biopsies obtained from four separate people and then immortalised *in vitro*. These four people are healthy controls 1

and 2, one LGMD2A (limb girdle muscular dystrophy type 2A) patient and one Pompe disease patient. Samples from two healthy controls were obtained via orthopaedic surgery. These myoblasts were cultured in the medium containing 15% FCS.

In the third batch, there were nine myotube cell lines. These myotubes were differentiated from myoblasts which were derived from the muscle biopsies obtained from patients and healthy controls and then immortalised *in vitro*. The myotubes were cultured in the medium containing 15% FCS.

The total number of the sample is shown in section 4.2.

2.1.3.3.   *Trophoblasts*

There were four sets of primary human trophoblasts isolated from four separate patients who had undergone an uncomplicated pregnancy and requested Cesarean section delivery. These trophoblasts were cultured *in vitro*. Four cytotrophoblast (CTB) and four syncytiotrophoblast (STB) samples were received in duplicate from Professor Gary F. Clark at University of Missouri. The STB was differentiated from the CTB.

Two batches of extravillous cytotrophoblast (evCTB) samples were provided by Dr. Sandra M. Blois at University Medicine Berlin, Germany.

2.2.   Methods

To acquire N-, O- and glycolipid glycans, samples were processed following an established protocol which has been summarized in Figure 2.1 (Jang-Lee et al., 2006; North et al., 2010; Jia et al., 2014; Parry et al., 2007).
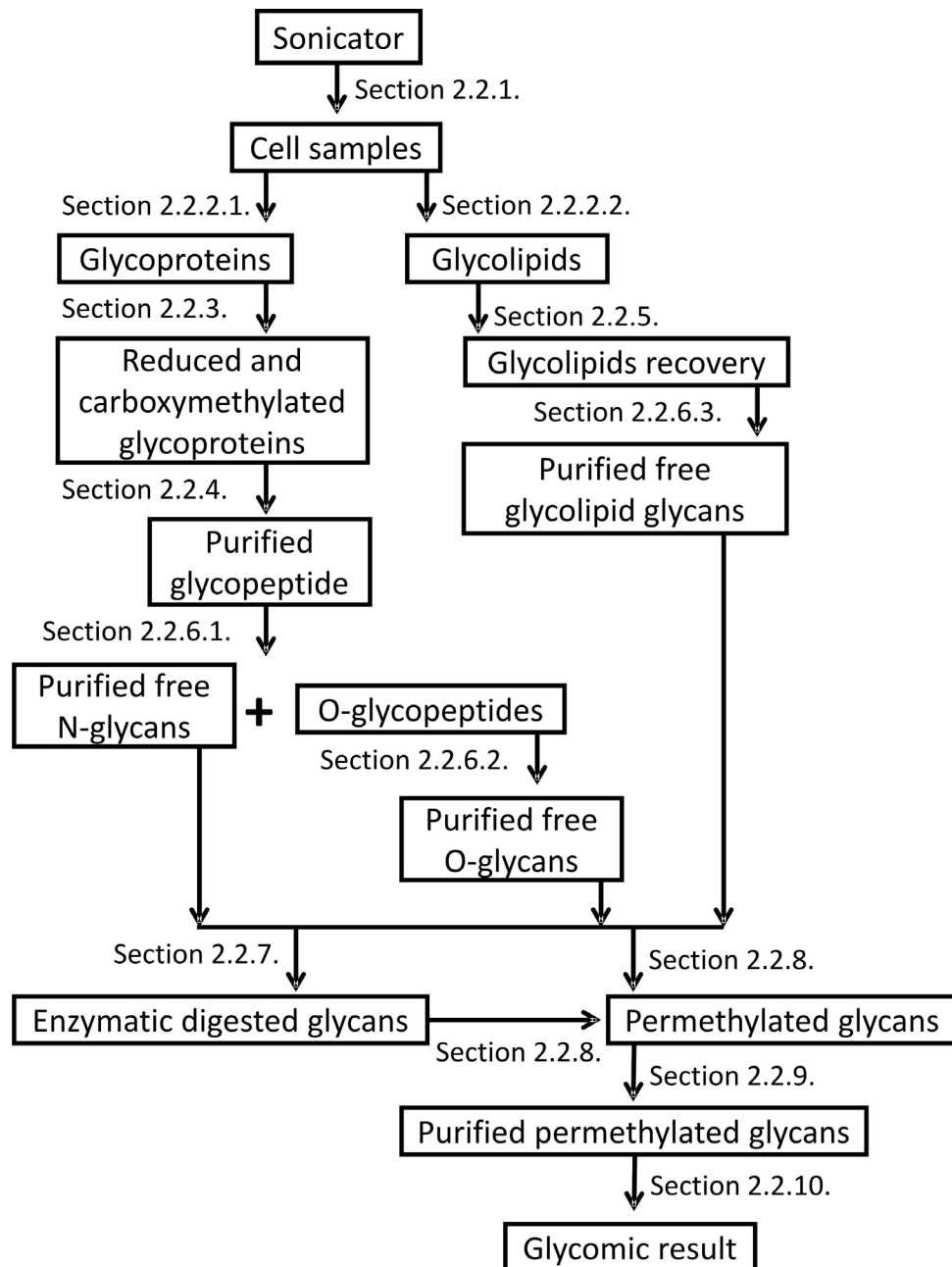
Figure 2.1 An overview of the glycomic approaches

2.2.1. Sonicator cleaning

It is important to clean the sonicator (VC130PB, Sonics & Materials, Inc.) that will be used in the homogenisation step as mass spectrometry is a very sensitive technique and any contaminant left on the sonicator could have an influence on the experiment result.

To clean the sonicator, firstly, the tip of the sonicator was dipped into ultra-pure water in a clean beaker and sonicated in continuous mode (10 seconds, 20 Amps). The first step was repeated with a solution containing 33.33% (volume percent) methanol, 33.33% formic acid and 33.33% ultra-pure water. After that the tip of the sonicator was immersed in ultra-pure water and kept in an ultrasonic bath (FS200, Decon) 10 minutes. The first step was repeated with following solutions: ultra-pure water, methanol, methanol chloroform mixture (v: v=1: 1), methanol, ultra-pure water (twice). Finally the sonicator was activated at 40 Amps in an empty Falcon tube for 1 minute.

2.2.2. Homogenisation

In order to extract the glycoproteins and glycolipids from cells, homogenisation is required to mechanically disrupt cells. However, because glycoproteins and glycolipids have different properties, different extraction approaches were employed.

2.2.2.1. *Glycoprotein extraction*

Lyophilized cell pellets were suspended in 1ml homogenization buffer (150 mM NaCl, 25 mM Tris, 5 mM EDTA and 1% CHAPS (volume percent), pH 7.4), sonicated in continuous mode for 10 seconds at 40 Amps. This was repeated 5 times pausing on ice for 15 seconds between each sonication.

Lysed sample was transferred into a dialysis cassette (3.5K MWCO, 0.5 - 3 ml, Thermo Scientific) and dialysed against the 4.5 L dialysis buffer (50 mM ammonia bicarbonate, pH 7.5) at 4 °C for 48 hours. Constant stirring was used during the dialysis process, the dialysis buffer was changed regularly (approximately every 12 hours). After dialysis, the sample was transferred into a 15 ml Falcon tube, covered with perforated Parafilm (Bemis) and lyophilized using a freeze dryer (Modulyod-230, Thermo Fisher).

2.2.2.2. Glycolipid extraction

Lyophilized cell pellets were suspended in 2 ml ice cold ultra-pure water and then transferred into a 15 ml Falcon tube. The sample was sonicated on ice in continuous mode (10 seconds, 40 Amps). This was repeated 2-3 times pausing on ice for 15 seconds between each sonication. Then the tube was taken into fume hood, in which 2.67 volumes of methanol (2.67 × volume of the ice cold ultra-pure water) was added and mixed fully. After that 1.33 volumes of chloroform (1.33 × volume of the water) was added and mixed fully. The mixture was centrifuged (IEC Centra CL3 centrifuge, Thermo) at 3000 rpm for 10 minutes (100 µl 0.6 M Tris buffer was added when the protein did not precipitate, and then the mixture was centrifuged again). The final supernatant (glycolipid fraction) was collected and its volume was measured. Then 0.173 volume of ultra-pure water (0.173 × volume of the supernatant) was added and mixed fully. The mixture was centrifuged (3000 rpm, 15 minutes). The mixture was divided into two layers which were collected separately: up layer is polar glycolipid; bottom layer is nonpolar glycolipid.

2.2.3. Reduction and carboxymethylation

Tris buffer (0.6 M, pH 8.5) was freshly made; pH was adjusted using acetic acid. The buffer was degased by bubbling gently nitrogen at the bottom of the buffer for 30 minutes.

The lyophilized sample from section 2.2.2.1 was suspended in 200 µl degased Tris buffer. Then 200 µl 10 mg/ml DTT (prepared from the degased Tris buffer) was added. The mixture was incubated (37 °C, 90 minutes). After the incubation the sample was centrifuged shortly. 200 µl 60 mg/ml IAA (prepared from the degased Tris buffer) was then added, this mixture was incubated at room temperature in dark (approximately 20 °C, 90 minutes). The reaction was terminated by transferring the sample into a dialysis cassette. The mixture was dialysed in 4.5 L dialysis buffer (see section 2.2.2.1) at 4 °C for 48 hours. Constant stirring was also used during the dialysis process; the dialysis buffer was changed regularly as mentioned previously (see section 2.2.2.1). After dialysis, the sample was transferred into a clean disposable glass tube (GPI 15-415, 13x100mm, Corning), covered with perforated Parafilm and lyophilized.

2.2.4. Tryptic digestion and glycopeptide purification

Ammonium bicarbonate (AMBIC) buffer (50 mM, pH 8.4) was freshly made; pH was adjusted using ammonia solution.

The lyophilized sample from section 2.2.3 was dissolved in 100 µl AMBIC buffer, then 200 µl 1 mg/ml trypsin (prepared from the AMBIC buffer) was added. The mixture was incubated (37 °C, 14-16 hours). After the incubation the sample was centrifuged shortly. The tryptic digestion was terminated by heating the sample (100 °C, 2 minutes). 1 drop of acetic acid was added to neutralize the sample. The sample was then preceded to the purification step.

A purification cartridge (Oasis HLB Plus, Waters) was attached to a 5 ml glass syringe (Samco) and then conditioned by eluting successively with 5 ml methanol, 5 ml 5% acetic acid (volume percent), 5 ml propan-1-ol, and 15 ml 5% acetic acid. The sample was loaded dropwise onto the cartridge. In order to avoid losing sample, the tube was washed 1-2 times using 5% acetic acid, the washing solution was also loaded onto the cartridge. The cartridge was washed using 20 ml 5% acetic acid, 4 ml 20% propan-1-ol solution (volume percent, 20% propan-1-ol in 80% 5% acetic acid), 4 ml 40% propan-1-ol solution, 4 ml 100% propan-1-ol. All propanol fractions were collected while the 5% acetic acid fraction was discarded as it contained hydrophilic contaminants. The volume of the fractions was reduced in a concentrator centrifuge, SpeedVac (SPD121P-230, Thermo) until these fractions could be combined into a single tube. Finally, the sample was covered with perforated Parafilm and lyophilized.

2.2.5. Polar glycolipid recovery

The step was carried out in the fume hood due to the use of chloroform. A purification cartridge (tC18, Sep-Pak, Waters) was attached to a 5 ml glass syringe (Samco) and then conditioned by eluting successively with 5 ml methanol, 5 ml methanol water mixture (v: v=1: 1), 5 ml methanol chloroform mixture (v: v=1: 1), and again 15 ml methanol water mixture. The polar glycolipid fraction from section 2.2.2.2 was loaded dropwise onto the cartridge.

The cartridge was washed using 15 ml methanol water mixture, 5 ml methanol and 5 ml methanol chloroform mixture. The methanol containing fractions were collected, combined and dried under nitrogen.

2.2.6. Release of glycans from glycoconjugates and purification of released glycans

As previously mentioned in sections 1.2.2.1, 1.2.2.2 and 1.2.3.1, N-, O- and glycolipid glycans are linked to proteins and lipids in different patterns, the approaches that employed to release these glycans are various.

All N-glycans were cleaved by peptide-N-glycosidase F (PNGase F). PNGase F is an endoglycosidase that cleaves the linkage between the N-glycan inner most GlcNAc and the amine group of the GlcNAc linked Asn residue on glycoproteins (Varki, 2009; Tarentino et al., 1985). This cleavage did not cause any change on the glycan portion but resulted in variation on the protein backbone: the Asn residue is substituted by an aspartic acid (Asp) residue.

All O-glycans were released by alkaline borohydride treatment (alkaline elimination) (Wada et al., 2010; Varki, 2009). Although O-glycosidase is also available, but it only cleaves some simple core 1 glycans, thus a chemical method, alkaline elimination, was used to release O-glycans. The reaction was performed under reducing environment which is helpful for preventing 'peeling reaction' of the released glycans via reducing the terminal GalNAc to its alditol form, thus it is also termed reductive elimination. This elimination was usually carried out after N-glycan release; otherwise it would cleave both N- and O-glycans, which would complicate subsequent data analysis.

The glycolipid glycans were released using ceramide glycanase (Recombinant endoglycoceramidase II). Like PNGase F, this enzyme is also an endo-enzyme which cuts in the chain between the glycan and the ceramide (see section 1.2.3.1) (Izu et al., 1997).

2.2.6.1.   Release of N-glycans and purification of released N-glycans

AMBIC buffer was freshly made (see section 2.2.4 for details). PNGase F cleaves N-glycan chains from glycoproteins/glycopeptides. However, it cannot cleave the glycan chain which carries an α1,3-linked core fucose residue at the innermost GlcNAc (Tretter et al., 1991; Parc et al., 2015).

The lyophilized sample from section 2.2.4 was dissolved in AMBIC buffer (200-300 μl, depends on the amount of the sample), then 4 U of PNGase F was added (the amount of the enzyme added is determined by the amount of the sample) and incubated (37 °C, 24 hours, another 4 U of PNGase F after the first 12-hour incubation). After incubation the sample was covered with perforated Parafilm and lyophilized. The lyophilized sample was proceeded to purification.

A classic C18 Sep-Pak cartridge (Waters) was attached to a 5 ml glass syringe and then conditioned as previously mentioned in section 2.2.4. The sample was loaded dropwise onto the cartridge. The cartridge was washed using 5 ml 5% acetic acid, followed by 4 ml 20% propan-1-ol solution, 4 ml 40% propan-1-ol solution, 4 ml 100% propan-1-ol. The 5% acetic acid fraction was collected as it contained released N-glycans. All propanol fractions were also collected as they contained the remaining peptides and O-glycopeptides. The volume of the fractions was reduced in SpeedVac until the fractions could be combined into a single tube. Finally, both 5% acetic acid fraction and propanol fraction were covered with perforated Parafilm and lyophilized.

2.2.6.2.   Release of O-glycans and purification of released O-glycans

KOH solution (50 ml, 0.1 M) was made, and it is used for KBH_4 solution (1 M) preparation.

The lyophilized propanol fraction from section 2.2.6.1 was dissolved in KBH_4 solution (400 μl), and then incubated (45 °C, 20–24 hours). After incubation, acetic acid was added dropwise until there was no bubbling. The sample was briefly centrifuged and proceeded to the next step.

A small amount of glass wool was filled at the narrow end of a Pasteur pipette, and the narrow end was inserted into a section of silicone tube, the internal bore diameter of the tube is 1–2 mm. A screw-adjustable switch was attached close to the tip of the tube. This Pasteur pipette was fixed by a retort stand. The pipette was filled with 5% acetic acid, and then the switch was opened to make the acid run out slowly. At the same time, the pipette was refilled with ion-exchange Dowex 50W-X8 beads which were used to remove peptides and cationic salts. The beads were washed using 5% acetic acid (20 ml). The switch was employed to control flow to keep the level of the acid above the bead level. The sample was loaded dropwise into the pipette. The acid should be controlled to pass slowly the pipette. The beads were washed using 2 × 3 ml 5% acetic acid. The two 3 ml acetic acid fractions were collected though usually the first 3 ml contains released O-glycans. The volume of the fractions was reduced in SpeedVac to approximately 1 ml, and then covered with perforated Parafilm and lyophilized. However, the borates could not be removed by beads, thus an extra step, coevaporation, was required: after lyophilisation 0.5 ml methanolic acetic acid (volume percent, 10% acetic acid in 90% methanol) was added to the lyophilized sample, the mixture was dried under nitrogen, and this was repeated 3-4 times.

2.2.6.3.   *Release of glycolipid glycans and purification of released glycolipid glycans*

Sodium acetate buffer (50 mM, pH 5.5) was freshly made, pH was adjusted using 5% acetic acid, sodium cholate was finally added to the buffer make its concentration as 0.2% (v/w).

The lyophilized sample from section 2.2.5 was resuspended in the sodium acetate buffer (200 μl). Then 25 mU ceramide glycanase was added to the sample and the mixture was incubated (37 °C, 24 hours). After this another 25 mU of the enzyme was added and the incubation was carried out for another 24 hours. After incubation, ultra-pure water was added to the sample to make the total volume 2 ml. Then 2 ml butanol was added. The mixture was centrifuged vigorously, the up layer was discarded. This step was repeated 1-2 times. Trace of butanol that left in the remaining liquid fraction was removed via drying under nitrogen.

A classic C18 Sep-Pak cartridge was attached to a 5 ml glass syringe and then conditioned by eluting successively with 5 ml methanol, 5 ml 5% acetic acid, 5 ml acetonitrile and 15 ml 5% acetic acid. The sample was loaded dropwise onto the cartridge, the cartridge was then

washed using 5 ml 5% acetic acid. The acetic acid fraction was collected using a clean glass culture tube.

This sample was further purified using Hypercarb column (Thermo Scientific) chromatography. The column was conditioned by eluting successively with 3 column volumes of 80% acetonitrile in 0.1% TFA (volume percent, the remaining 19.9% is ultra-pure water) and 3 column volumes of ultra-pure water. The sample was loaded dropwise onto the cartridge, the cartridge was then washed using 3 column volumes of ultra-pure water and 2 column volumes of 25% acetonitrile in 0.05% TFA (volume percent, the remaining 74.95% is ultra-pure water). The acetonitrile fraction was collected. The volume of the fractions was reduced in SpeedVac until the fraction volume was approximately 1 ml. Finally, the sample was covered with perforated Parafilm and lyophilized.

2.2.7.   Other enzymatic digestions and digested glycan purification

Released glycans from glycoproteins and glycolipids can be digested by enzymes, such as sialidase S, sialidase A, β1,4-galactosyltransferase and endo-β-galactosidase. These digestions can provide extra structural information.

2.2.7.1.   *Sialidase S digestion*

Sialidase S is an enzyme which specifically removes non-reducing terminal unbranched α2,3 linked sialic acid from glycoconjugates (Corfield et al., 1983). Sodium acetate buffer was freshly made as mentioned in section 2.2.6.3. 1 unit sialidase S powder was dissolved in the buffer (200 μl) to make 5mU/μl enzyme solution.

The sample from previous was dissolved in the buffer (200 μl). 20 μl of enzyme solution was added to the sample, then the sample was incubated (37 °C, 24 hours, another 20 μl of the enzyme solution was added to the sample after the first 12 hours). After incubation, the sample was briefly centrifuged and proceeded to the purification step.

2.2.7.2.   Sialidase A digestion

Sialidase A is an enzyme which can specifically cleave all non-reducing terminal sialic acid from glycoconjugates (Ohta et al., 1989). 1 unit sialidase A powder was made into solution following the procedure in section 2.2.7.1.

The sample from previous section was treated with sialidase A following the procedure that described in section 2.2.7.1.

2.2.7.3.   β1, 4-galactosyltransferase reaction

β1,4-galactosyltransferase is the enzyme which can transfer a Gal from a UDP-Gal to a GlcNac producing a disaccharide unit, Galβ1,4GlcNac. However, if a GlcNac is bisected (see section 1.2.2.1), it will not be modified by this enzyme (Qasba et al., 2008; Narasimhan et al., 1985). 3-(N-Morpholino) - propanesulfonic acid (MOPS) solution (50 mM, pH 7.4, contains 45 uM UDP-Gal) was freshly made, pH was adjusted using ammonia, and then manganese chloride 4-hydrate was added to make its final concentration as 20 mM. 1 unit β1,4-galactosyltransferase was dissolved in the solution (200 µl).

The sample from previous section was dissolved in the enzyme solution (50 µl), then the sample was incubated (37 °C, 24 hours, another 50 µl of the enzyme solution was added to the sample after first 12 hours). The sample was briefly centrifuged and proceeded to the next step.

2.2.7.4.   Endo- β-galactosidase digestion

Endo-β-galactosidase is an enzyme which hydrolyses internal β1,4 galactosidic linkage in this favoured repeating unit [GlcNAcβ1,3Galβ1,4]$_n$ (Scudder et al., 1983). Sodium acetate buffer (100 mM, pH 5.8) was freshly made; pH was adjusted using 5% acetic acid. 0.1 unit endo-β-galactosidase powder was dissolved in the buffer (100 µl).

The sample from previous section was dissolved in the sodium acetate buffer (150 µl). 20 µl of enzyme solution was added to the sample, then the sample was incubated (37 °C, 48 hours, another 20 µl of the enzyme solution was added to the sample after first 24 hours). The sample was briefly centrifuged and proceeded to the separation step.

2.2.7.5.   *Digested glycan purification*

A purification classic C18 Sep-Pak cartridge (Waters) was attached to a 5 ml glass syringe and then conditioned as described in section 2.2.4. The sample was loaded dropwise onto the cartridge. The cartridge was washed using 5 ml 5% acetic acid. The 5% acetic acid fraction was collected as it contained glycans required. The volume of the fractions was reduced to about 1 ml by SpeedVac. Finally, the fraction was covered with perforated Parafilm and lyophilized.

2.2.8.   Permethylation of released glycans

3-5 sodium hydroxide pellets and 3 ml dimethyl sulfoxide (DMSO) were added in a mortar and ground with a pestle. The pellet should be ground into slurry. The slurry (about 1 ml) was added to the glass tube which contained lyophilized sample from previous section. Then methyl iodide (about 0.6 ml) was added into the tube. The tube was capped and fixed on an automatic vortexer (VX-2500, VWR) and vortexed (15-20 minutes, 20 °C). The reaction was terminated by addition of ultra-pure water dropwise until there was no bubbling. Then 1 ml chloroform was added, ultra-pure water was added to make the final volume of the mixture 5 ml. The mixture was centrifuged (2500 rpm, 30 seconds), the up layer was discarded. This step was repeated 4 times. The bottom layer was chloroform layer which contained permethylated glycans. It was dried under a gentle stream of nitrogen.

Figure 2.2 shows the reducing end (OR), the non-reducing end (R) and the method of how to calculate the mass of a permethylated glycan.
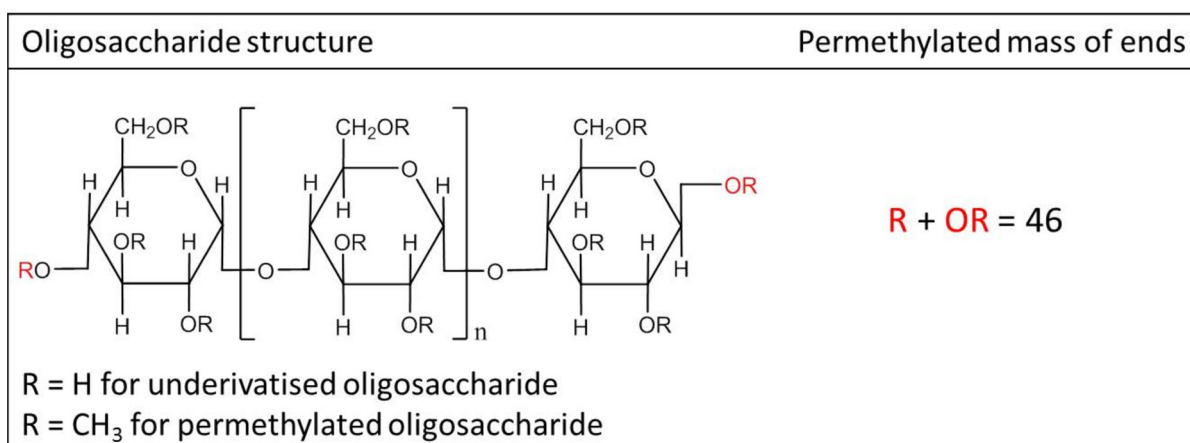
Figure 2.2 The method used to calculate the mass of a permethylated glycan residue

2.2.9. Purification of permethylated glycans

1.5 ml, 3.5 ml, 5 ml and 7.5 ml acetonitrile was mixed with 8.5 ml, 6.5 ml, 5 ml and 2.5 ml ultra-pure water respectively to make 10 ml 15%, 35%, 50% and 75% acetonitrile solution.

A classic C18 Sep-Pak cartridge (Waters) was attached to a 5 ml glass syringe and then conditioned by eluting successively with 5 ml methanol, 5 ml ultra-pure water, 5 ml acetonitrile, and 15 ml ultra-pure water. The sample from section 2.2.8 was dissolved in methanol/ultra-pure water (200 µl, v:v=1:1) and then loaded dropwise onto the cartridge. The cartridge was washed using 5 ml ultra-pure water following by 3 ml 15%, 35%, 50% and 75% acetonitrile solution. All acetonitrile fractions were collected. The volume of the fractions was reduced to about 1 ml by SpeedVac. Finally, the fractions were covered by perforated Parafilm and lyophilized.

2.2.10. Mass spectrometry

The mass spectrometric techniques used are matrix assisted laser desorption ionization time-of-flight (MALDI-TOF), matrix assisted laser desorption ionization time-of-flight/time-of-flight (MALDI-TOF/TOF) and gas chromatography mass spectrometry (GC-MS).

2.2.10.1. MS data acquisition and analysis

The MS matrix, 2,5-dihydroxybenzoic acid (DHB) (20 mg/ml) in methanol/water mixture (v:v=7:3), for permethylated glycans was freshly made. The MS matrix, α-cyano-4-hydroxycinnamic acid (10 mg/ml) in acetonitrile/TFA/water mixture (v:v:v=500:1:499), for calibration standard was freshly made.

Purified permethylated glycans from section 2.2.9 was dissolved in methanol (10 µl). 1 µl of the sample was combined with 1 µl of DHB matrix and then loaded onto a metal plate. In addition, prepared calibration standard (AB Sciex) mixed with calibration matrix was also spotted on the plate. After the sample and calibration standard were dried, the plate was put into a Voyager-DE STR MALDI workstation (Applied Biosystems) mass spectrometer, via which the calibration was carried out and the MS data could be acquired. The Voyager was run in the reflectron positive ion mode and its accelerating voltage was set as 20kV.

Obtained MS data was processed using Data Explorer Software version 4.9 (Applied Biosystems). Generally, the MS spectrum was baseline corrected with the default setting and then smoothed with the noise filter at a correlation factor of 0.7. The processed spectrum was annotated using a glycoinformatic tool, GlycoWorkBench (Ceroni et al., 2008) and a graphic design software, CorelDraw X3.

2.2.10.2. MS/MS data acquisition and analysis

The MS/MS matrix, 3,4-diamino-benzophenone (DABP) (10 mg/ml) in acetonitrile/water mixture (v:v=7:3), for permethylated glycans was freshly made. The MS/MS matrix for calibration standard is the same as the MS matrix for calibration standard described in section 2.2.10.1.

Previous dissolved sample from section 2.2.10.1 was dried and then redissolved in methanol (10 µl), 1 µl of the sample was combined with 1 µl of DABP matrix and then loaded onto a metal target plate. In addition, prepared calibration standard (AB Sciex) mixed with calibration matrix was also spotted on the plate. After the sample and calibration standard were dried, the plate was loaded into a 4800 MALDI-TOF/TOF mass spectrometer (AB SCIEX), via which the calibration was carried out. MS spectrum was firstly obtained, and

then peaks of interest were selected for MS/MS analyses and the MS/MS data could be acquired. The 4800 instrument was also run in the reflectron positive ion mode. The collision energy was set to 1 kV with argon as the collision gas.

Obtained MS/MS data was also processed using the same method described in section 2.2.10.1.

2.2.10.3. *GC-MS*

GC-MS requires the analyte to be volatile but the permethylated glycans are not volatile, therefore, permethylated glycans cannot be analysed directly. Further treatment of these glycans is required.

2.2.10.3.1. GC-MS sample preparation

TFA solution (2 M) was made freshly using ultra-pure water. $NaBD_4$ solution (10 mg/ml) was freshly made using 2 M ammonia.

TFA solution (200 µl) was added to the sample left in previous step and incubated (121 °C, 2 hours). After the incubation, the sample cooled down to the room temperature, briefly centrifuged and then dried under a gentle stream of nitrogen. $NaBD_4$ solution (200 µl) was added to the dried sample and incubated (room temperature, 2 hours). The incubation was terminated by adding acetic acid dropwise until there was no bubbling. The sample was redried under nitrogen. Because borates were introduced, coevaporation was required (see section 2.2.6.2). After coevaporation, acetic anhydride (200 µl) was added to the sample and then the mixture was incubated for acetylation (100 °C, 1 hour). Then the sample was briefly centrifuged and redried under nitrogen. 1 ml chloroform was added to the dry sample and then ultra-pure water was added to make the total volume 5 ml, the mixture was vortexed and then allowed it to separate into two layers, the up layer was discarded. This washing step was repeated 4 times. The bottom chloroform fraction was finally dried under a gentle stream of nitrogen. The final partially methylated alditol acetates (PMAA) were ready for analysis.

2.2.10.3.2. GC-MS data acquisition and analysis

A blank (usually hexane) was injected into RTX-5MS column (Restek Corp.) which was fitted in Perkin Elmer Clarus 500 GC-MS. This was performed to make sure the column was clean and the instrument was in good condition (could be used in the experiment). A glycan standard was firstly run and then the sample; retention times of the standard would be compared with the sample and would help us determine the compound of the sample. Sample from section 2.2.10.3.1 was usually dissolved in hexane (usually 50 µl, depends on the amount of the sample), and then 2 µl of dissolved sample was injected. All injections were carried out after the oven temperature reached 60 °C. Running a blank was required before every sample running.

A list of retention times and characteristic ions were prepared for comparison. Obtained GC-MS data, the gas chromatographic profiles and electron ionization spectra were processed using TurboMass version 4.5.0 (Perkin Elmer Instruments).

Chapter 3

Comparative glycomic profiling of HeLa cells and normal human dermal fibroblast (NHDF)

3. Comparative glycomic profiling of HeLa cells and normal human dermal fibroblast (NHDF)

3.1. Introduction to the project

This project is in collaboration with Professor Els Van Damme and Miss Chenjing Shang at Ghent University.

In addition to performing glycan array analyses of the five SNAs (Shang and Van Damme, 2014) mentioned in section 1.6.3, Van Damme and Shang have carried out cytotoxicity assay of these five SNAs using HeLa and NHDF cells, and found that all the SNAs are more toxic towards HeLa than NHDF cells. They have also performed *in vivo* luciferase (internalization) assay of SNAs using HeLa cells and found that the SNAs could get into the cell (Shang et al., 2015). All these made us hypothesize that HeLa and NHDF cells could express glycans that the SNAs can bind to and the patterns of the glycans in these cells would be different leading to the different cytotoxicities, and thus glycomic analysis of these two types of cells were carried out to check the hypothesis. Specifically, MS and MS/MS analyses provided general structural information of the glycans; the result of the glycan array analyses suggested the determination of sialic acid linkages in the glycans as the five SNAs favoured different glycan motifs (see section 1.6.3); quantification of nonsialylated (gal-terminated) and sialylated glycans may provide a possible explanation for the result of the cytotoxicity assay of these five SNAs.

The glycome patterns on both glycoproteins and glycolipids of HeLa and NHDF cells (see section 3.2) were analysed using MS to test the hypothesis.

3.2. Sample details and sample processing

Two batches of HeLa and NHDF cells were cutured in the same media (Shang et al., 2015) but on different days. The cells were sent to our laboratory once being harvested. The detail of two batches of NHDF and HeLa cells is shown in Table 3.1.

Table 3.1 Detail of NHDF and HeLa cells received from Ghent University

Cell type	Cell number
First batch	
NHDF	Approximately 4 million
HeLa cells	Approximately 7.5 million
NHDF	Approximately 7.5 million
HeLa cells	Approximately 10 million
Second batch	
Cell type	Cell number
NHDF	Approximately 4 million
HeLa cells	Approximately 21 million
NHDF	Approximately 7.3 million
HeLa cells	Approximately 30 million

All samples were analysed using glycomic methodologies which have been described in section 2.2.

3.3. Results

3.3.1. HeLa and NHDF N-glycans

3.3.1.1. MALDI-TOF MS analysis of the N-glycans of HeLa and NHDF cells

Duplicate analyses of N-glycans were carried out to check the reproducibility (Figure 3.1) and whether combining the duplicate for the more detailed structural research would be valid.
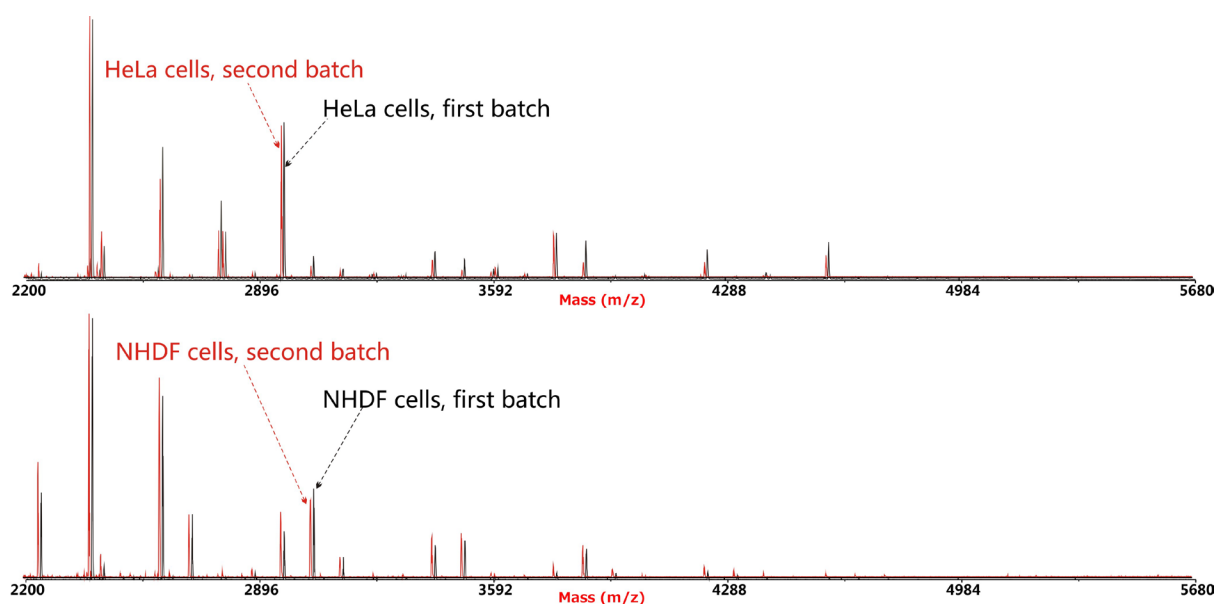
Figure 3.1 MALDI-TOF MS spectra of permethylated N-glycans (m/z 2200-5680) from two batches of HeLa cells (top) and two batches of NHDF cells (bottom)

The MS spectra of the cells from the first batch are black; the MS spectra of the cells from the second batch are red. The black spectra are overlaid on the red spectra with an offset of approximately 8 Da. In HeLa cells, the average peak variation between the spectra from different batches is approximately 33%. In NHDF cells, the average peak variation between the spectra from different batches is approximately 19%. The result shows that there is no significant difference in the N-glycan profile of biological replicates.

High quality MALDI data were obtained for the N-glycans from the HeLa and NHDF cells. Here 'high quality' means: 1. The spectrum does not show 14 dalton smaller underpermethylated peaks which could be due to the fact that the hydrogen of one of the hydroxyl groups in the glycan is not replaced by a methyl group (Morris et al., 1996a); 2. When zoom in the spectrum, the isotope peaks observed in the spectrum can be clearly distinguished; 3. The signal to noise ratio (S/N) of the spectrum is acceptable, which means the signal peaks of analyte ions do not merge with the baseline ("noise") and can be distinguished. S/N is a measure of a peak signal intensity relative to the "noise" level (Williams et al., 2003; Krutchinsky and Chait, 2002). Representative MALDI-TOF MS spectra of their N-glycans are shown in Figure 3.2 and Figure 3.3. Molecular ions were observed up to approximately m/z 5700 indicating that very good sensitivity was being achieved. The spectra show that high mannose and complex glycans are present in both cell types. Non-core fucosylated glycans, such as m/z 2792 and 3603, are more evident in HeLa cells. The dominant complex N-glycans in both HeLa and NHDF are sialylated with NeuAc

or are terminated with uncapped Gal. The result of a more detailed comparison of the glycan profiles of these two cell types has been summarized in section 3.3.1.7.
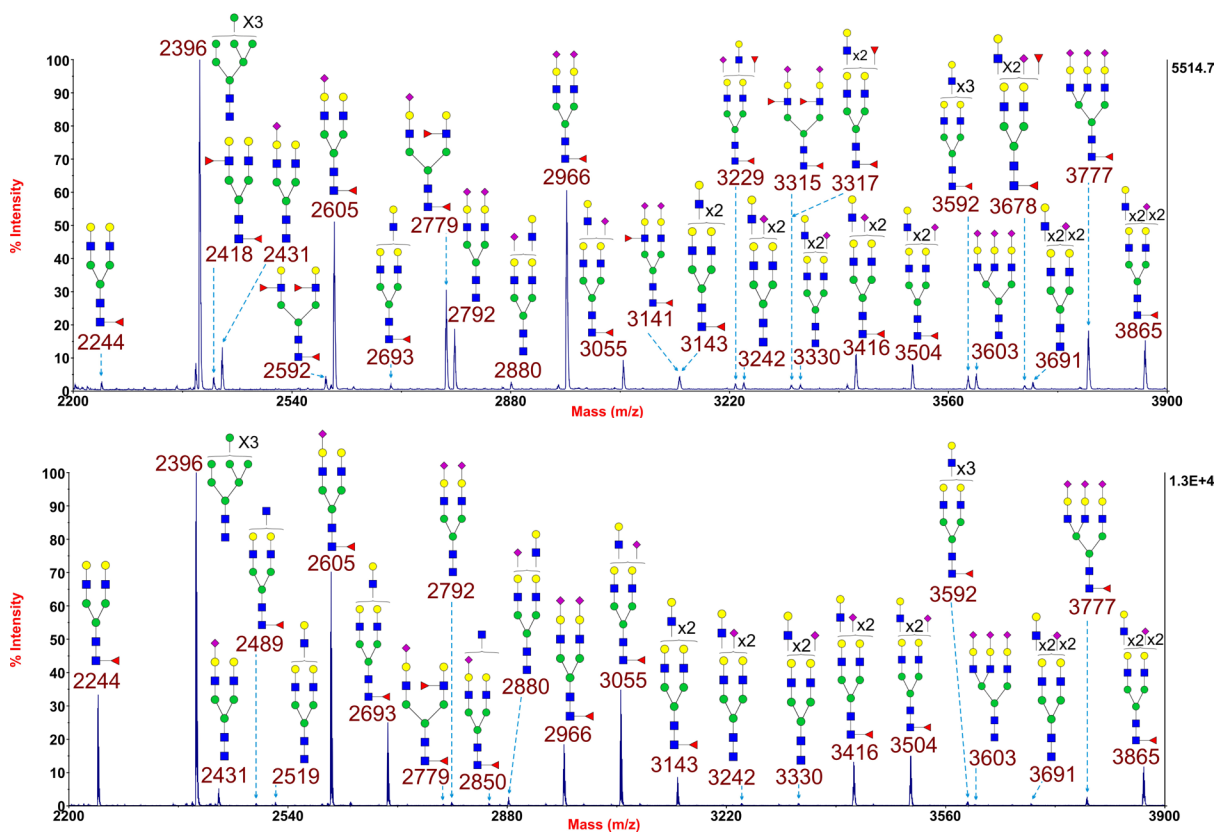
Figure 3.2 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2200-3900) from HeLa (top) and NHDF cells (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^{+}$. Putative structures are based on the molecular weight, N-glycan biosynthesis pathway and MS/MS data. Due to the presence of heterogeneous multiantennary structures with extended LacNAc (Gal-GlcNAc) repeats, the annotations are simplified throughout by using biantennary structures with the extensions or sugars listed outside a bracket. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
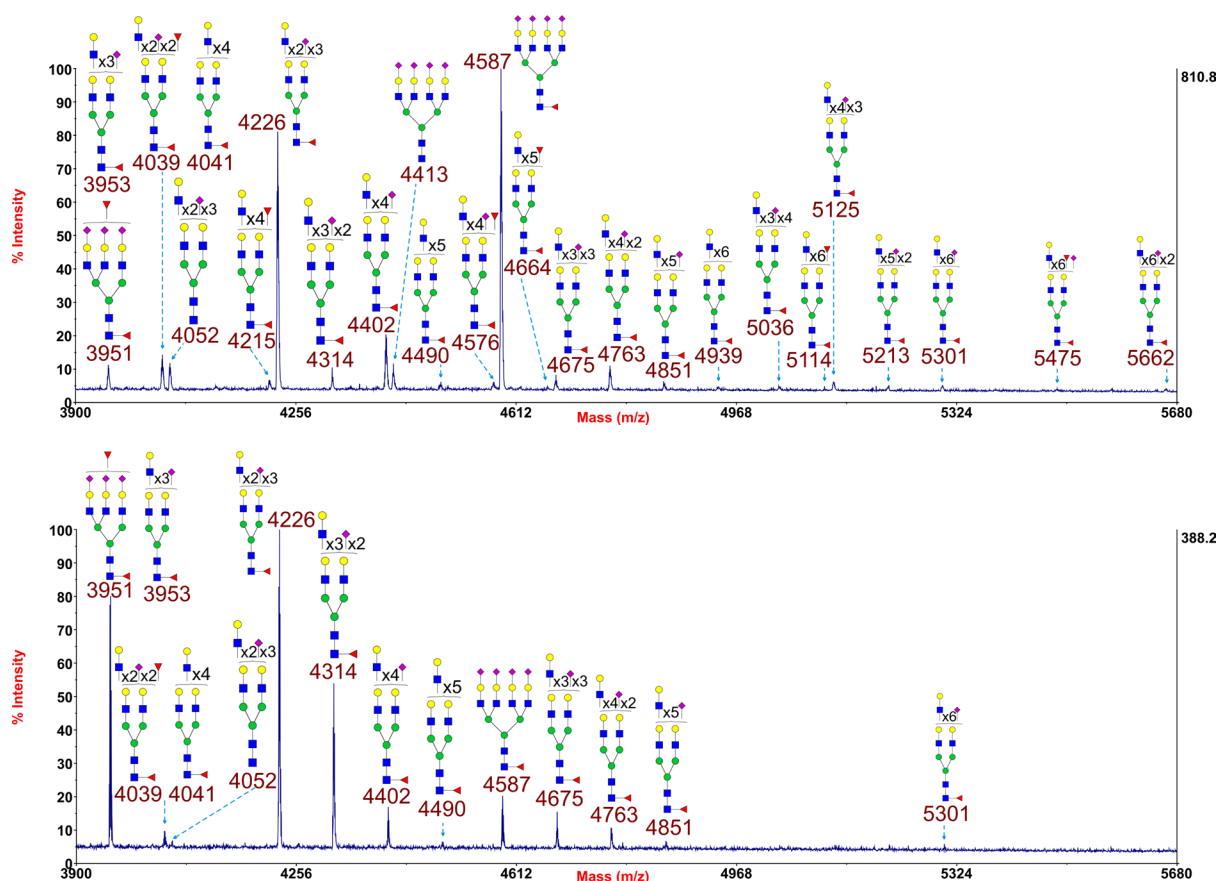
Figure 3.3 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 3900-5680) from HeLa (top) and NHDF cells (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

3.3.1.2. *MALDI-TOF/TOF MS/MS analyses of the N-glycan at m/z 3143 of HeLa and NHDF cells*

Peaks of interest were chosen for further characterization via MALDI-TOF/TOF MS/MS analysis. This analysis was used to distinguish isobaric glycans with different sequences as well as glycans of similar masses whose isotope clusters were overlapping.

Here the MS/MS data for the molecular ion clusters centred at m/z 3142 in two cells are described as an example of the latter. In HeLa cells the isotope peaks of the glycan(s) at m/z 3143 span a wider mass range than predicted for a single glycan composition, while the isotope pattern in NHDF is more akin to that expected for a glycan at that molecular weight (Figure 3.4). This implies that this peak cluster may include a LacNAc containing glycan as

well as a potential sialyl-Lewis X containing glycan which differ by 2 Da. This situation has been reported in dendritic cells and neutrophils (Babu et al., 2009; Bax et al., 2007).
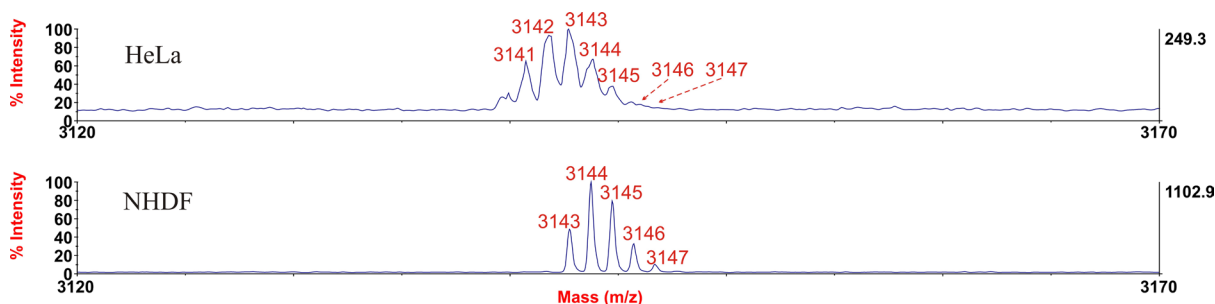
Figure 3.4 Isotope peak cluster is wider in HeLa cells than that in NHDF

The isotope peaks of m/z 3141 and 3143 are overlapping in spectra from HeLa cells, as a result of which the peak cluster is wider than in the NHDF data.

In order to confirm the presence of potential sialyl-Lewis X and LacNAc and their relative abundance, MS/MS analyses were carried out for the two samples. As shown in Figure 3.5 (top) the fragment ions at m/z 2679, 2230 and 1781 correspond to loss of LacNac, LacNac2 and LacNac3 respectively, and their concurrent ions are also observed at m/z 486 and 1385. This demonstrates the presence of oligo-LacNAc containing glycans, in accordance with the isotope pattern of the molecular ion cluster (Figure 3.4). The third most abundant fragment ion at m/z 2142 and its concurrent ion at m/z 1021 are attributable to a sialylated, fucosylated glycan. Fuc is confirmed to be 3-linked as the signal at m/z 2934 corresponds to the loss of a Fuc from the C3 position of GlcNAc via beta elimination. These provide evidence of the presence of potential sialyl-Lewis X. The abundance of the fragment ions corresponding to potential sialyl-Lewis X containing glycan suggests that it is minor compared to the LacNAc containing glycan in HeLa cells. In contrast to the HeLa data, MS/MS of m/z 3142 from NHDF cells showed fragment ions that correspond solely to oligo-LacNAc antennae (Figure 3.5 bottom), which is consistent with its isotope pattern.

Figure 3.5 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated N-glycan peak centred at m/z 3142 in HeLa cells (top) and NHDF cells (bottom)

Assignments of the fragment ions are indicated on the cartoons and on the spectra the horizontal red arrows show antennae losses whilst antennae-derived fragment ions are annotated with their sequences. The number indicated above the peak in the spectra is the m/z value of the fragment ion (resulting ion) that has been detected by the mass spectrometer. Data were acquired in the form of $[M+Na]^+$ ions. The glycan molecule and its corresponding fragment ions are labelled in the same colour (red or black). Due to the presence of heterogeneous multiantennary structures with extended LacNAc repeats, the annotations are simplified throughout by using biantennary structures with the extensions listed outside a bracket. ■ GlcNAc, ● Man, ● Gal, ▲ Fuc, ◆ NeuAc.

110

3.3.1.3. *Sialidase S digestion of the N-glycans of HeLa and NHDF cells*

It is necessary to know the linkage of the NeuAc because NeuAc linkage differences might cause the carbohydrate binding differences. To determine the NeuAc linkages, the glycans were digested with sialidase S and sialidase A, which cleave α2,3 linked and all non-reducing terminal NeuAc residues, respectively.

Digestion of HeLa and NHDF cells with sialidase S resulted in partial desialylation, and with sialylated N-glycans at m/z 2431, 2605, 3055 and 3504 being observed (Figure 3.6 and Figure 3.7). The comparison of Figure 3.2 (top) and Figure 3.3 (top) to Figure 3.6 (top) and Figure 3.7 (top) showed that the relative abundances of many sialylated glycans in HeLa cells were decreased, the most obvious decrease was observed in the glycans at m/z 2966, 3416, 3777, 3865, 4226 and 4587 which completely disappeared; simultaneously, the relative abundance of nonsialylated glycans was increased, the most obvious increase being observed in the glycans at m/z 2244, 2693 and 3143 which increased approximately 1004.9%, 1937.6% and 2152.5% respectively. Similarly, the comparison of Figure 3.2 (bottom) and Figure 3.3 (bottom) to Figure 3.6 (bottom) and Figure 3.7 (bottom) also showed that the relative abundance of many sialylated glycans decreased in NHDF cells, the most obvious decreases were observed in the glycans at m/z 2966, 3416, 3865, 4226 and 4587 which completely disappeared; simultaneously, the relative abundance of nonsialylated glycans increased, the most obvious increases were observed in the glycans at m/z 3143, 4041 and 4490 which were increased approximately 2080.7%, 2617.0% and 745.1% respectively. In addition, sialyl Lewis X disappeared after the desialylation (see section 3.3.1.4 for evidence). This indicated that both HeLa and NHDF contain a mixture of α2,3 and α2,6 linked NeuAc. The ratio of desialylated glycans in HeLa cells is slightly higher than that in NHDF cells (see section 3.3.1.7 for details).

Figure 3.6 Annotated MALDI-TOF MS spectra of permethylated sialidase S treated N-glycans (m/z 2200-3900) from HeLa (top) and NHDF cells (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]⁺. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
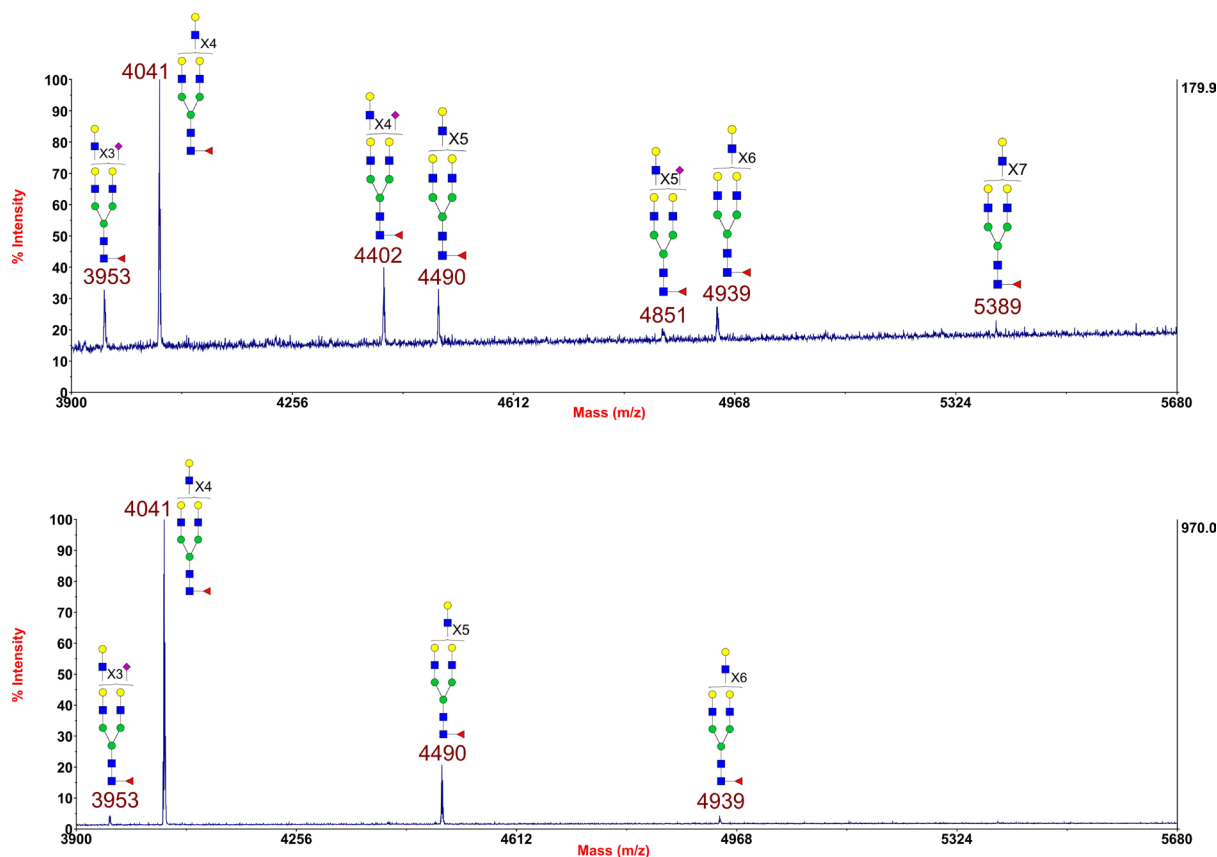
Figure 3.7 Annotated MALDI-TOF MS spectra of permethylated sialidase S treated N-glycans (m/z 3900-5680) from HeLa (top) and NHDF cells (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

3.3.1.4. *MALDI-TOF/TOF MS/MS analysis of glycan cluster at m/z 3143 of HeLa cells after sialidase S digestion*

After sialidase S digestion, the isotope peak of the glycan at m/z 3143 in HeLa cells is akin to that predicted. This implies that this peak only includes one glycan, probably the LacNAc containing glycan. In order to confirm this, MS/MS analysis was carried out. The presence of fragment ions at m/z 2230, 1781 and 1385 (Figure 3.8) demonstrated the presence of an oligo-LacNAc containing glycan only.
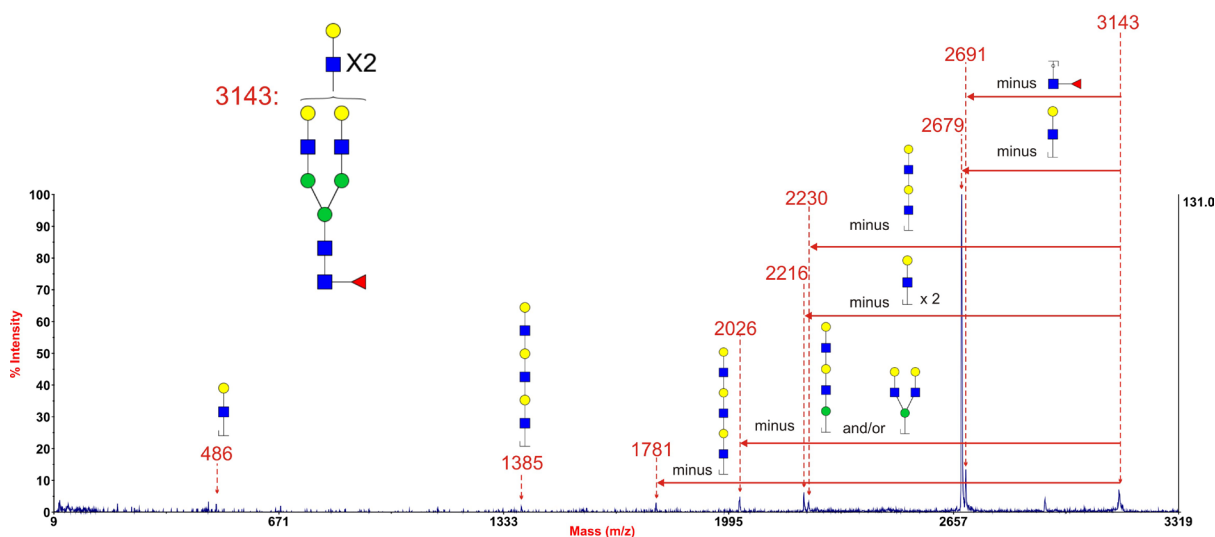
Figure 3.8 Annotated MALDI-TOF/TOF MS/MS spectrum of permethylated sialidase S treated N-glycan at m/z 3143 in HeLa cells

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

3.3.1.5.   *Sialidase A digestion of the N-glycans of HeLa cells and NHDF*

The glycan profiles are similar between the two cell types after sialidase A digestion. Treatment with sialidase A resulted in a complete desialylation of the N-glycans (Figure 3.9 and Figure 3.10), for instance, the sialyl Lewis X containing glycan at m/z 3141 in Figure 3.2 (top) became the glycan at m/z 2418 in Figure 3.9 (top). LacNAc or polyLacNAc containing glycans, such as m/z 2244, 2693 and 3143, were the main products of the enzymatic reaction. With the disappearance of sialylated N-glycans, glycans at 2244, 2693 and 3143 increased approximately 2309.5%, 3119.7% and 2208.7% in HeLa cells respectively, 100.9%, 145.1% and 1632.2% in NHDF cells respectively. In addition, some novel nonsialylated N-glycans which were not observed in Figure 3.2 were shown in Figure 3.9. For instance, glycans at m/z 2867 in HeLa cells and glycan at m/z 2418 in NHDF cells were only observed after sialidase A digestion. All these glycans (m/z 2418, 2592, 2867, 3317, 3766, 4215, 4664 and 5114) possessed more than one Fuc. Within the displayed mass range, the completely desialylated form of the glycan peaks annotated in Figure 3.2 and Figure 3.3 could all be found in Figure 3.9 and Figure 3.10.

These confirmed that in both HeLa and NHDF cells the complex glycans contained terminal

Gal and NeuAc as well as fucosylated antennae.
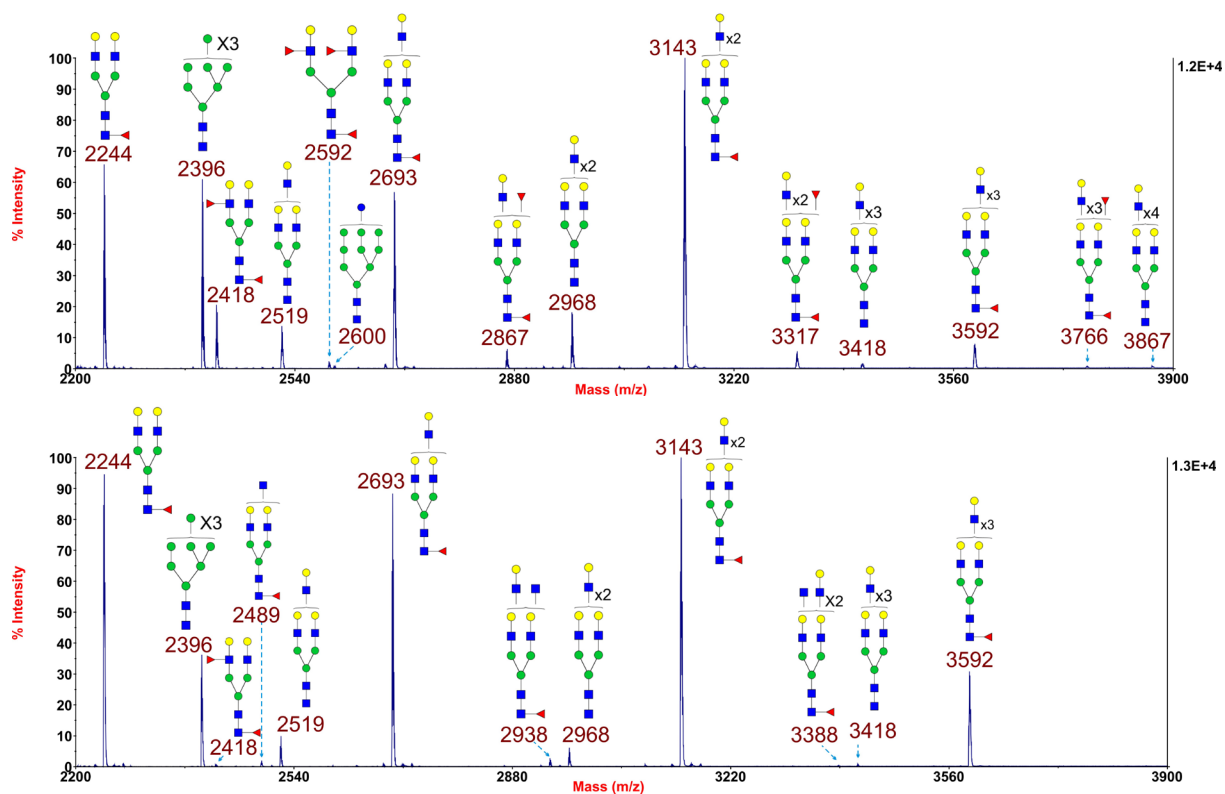
Figure 3.9 Annotated MALDI-TOF MS spectra of permethylated sialidase A treated N-glycans (m/z 2200-3900) from HeLa (top) and NHDF cells (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
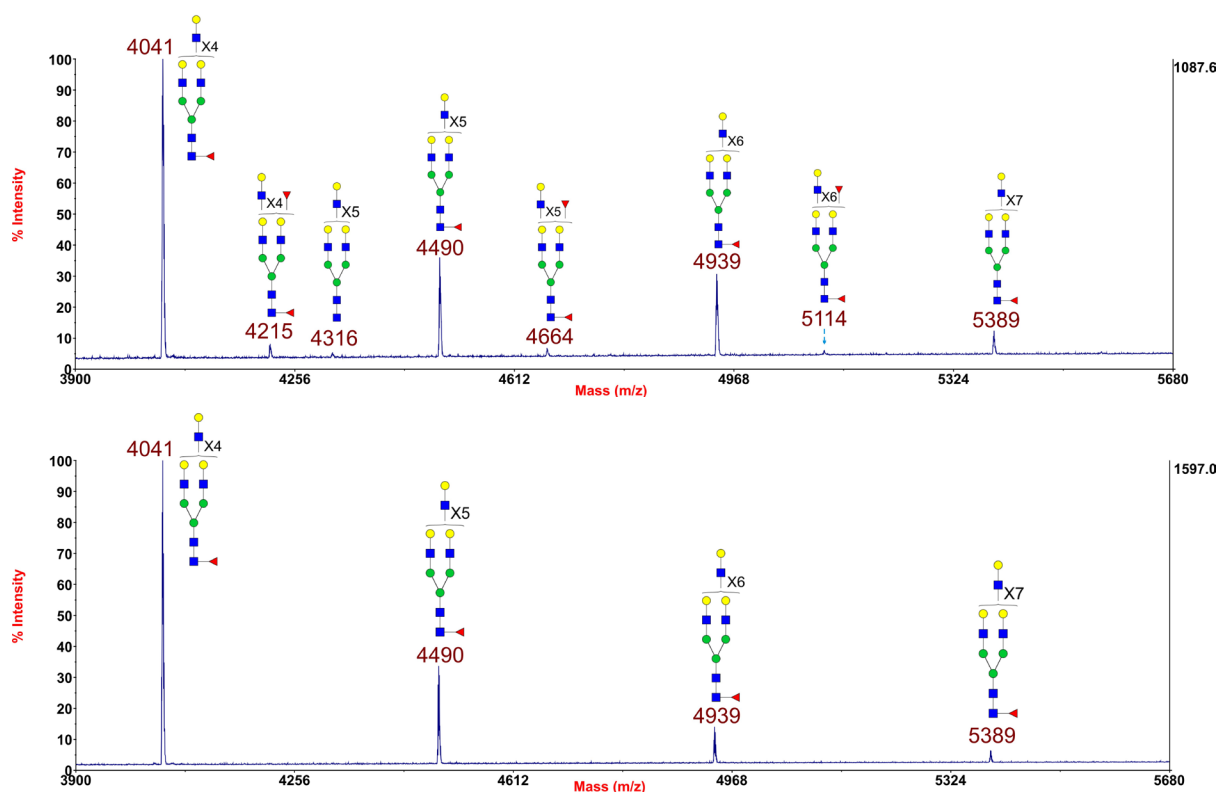
Figure 3.10 Annotated MALDI-TOF MS spectra of permethylated sialidase A treated N-glycans (m/z 3900-5680) from HeLa (top) and NHDF cells (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]⁺. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

3.3.1.6. *Quantification of the N-glycans*

Since terminal Gal and NeuAc play important roles in the carbohydrate binding, it is necessary to quantify the Gal and NeuAc and then make a comparison. Quantification of terminal Gal and terminal NeuAc on glycan antennae of each cell type was carried out via comparing the relative abundance of LacNAc (Gal-GlcNAc) antenna and sialylated LacNAc antenna.

The following figure shows the quantification method. According to the Figure 3.11A, the glycan at m/z 3055 is a mixture of biantennary and triantennary. The ratio of the two structures in the mixture was calculated by dividing the relative intensity sum of the corresponding characteristic ions. The assumption was made that the probabilities of losing one LacNAc and one sialylated LacNAc in biantennary structure are 50% and 50%

116

respectively, and the probabilities of losing one LacNAc and one sialylated LacNAc in triantennary structure are 67% and 33% respectively. Therefore the expected values of losing one LacNAc and one sialylated LacNAc are 0.51 and 0.49 respectively. Since the relative ion count of the glycan at m/z 3055 is 4452.7 (Figure 3.11B); it is possible to calculate the individual ion abundance of the LacNAc and sialylated LacNAc.

A
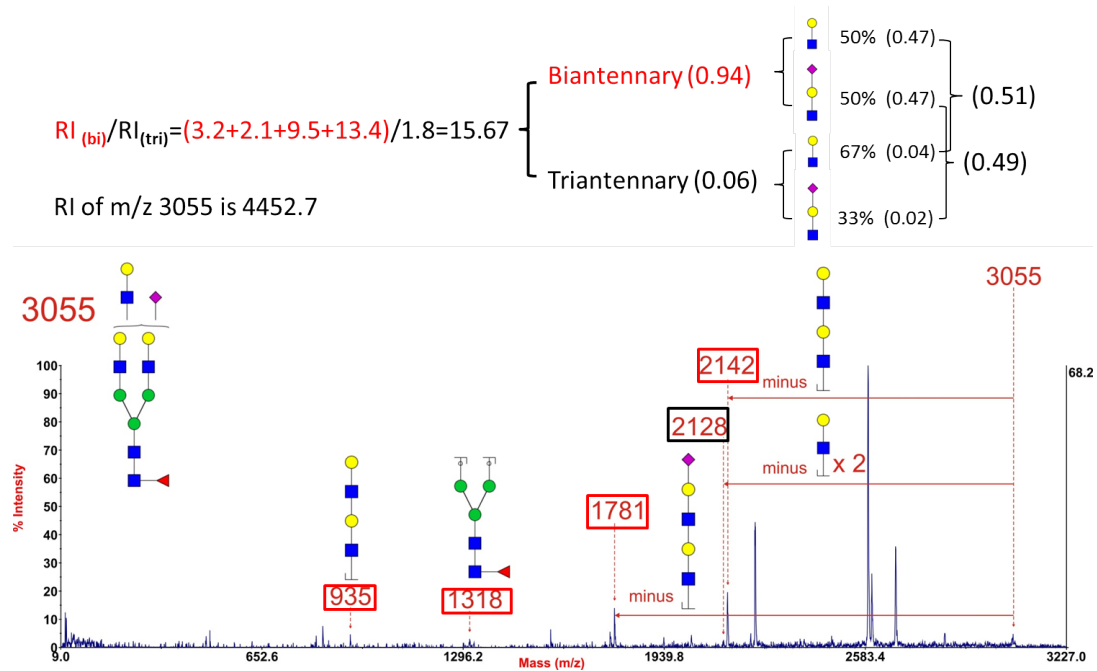
B
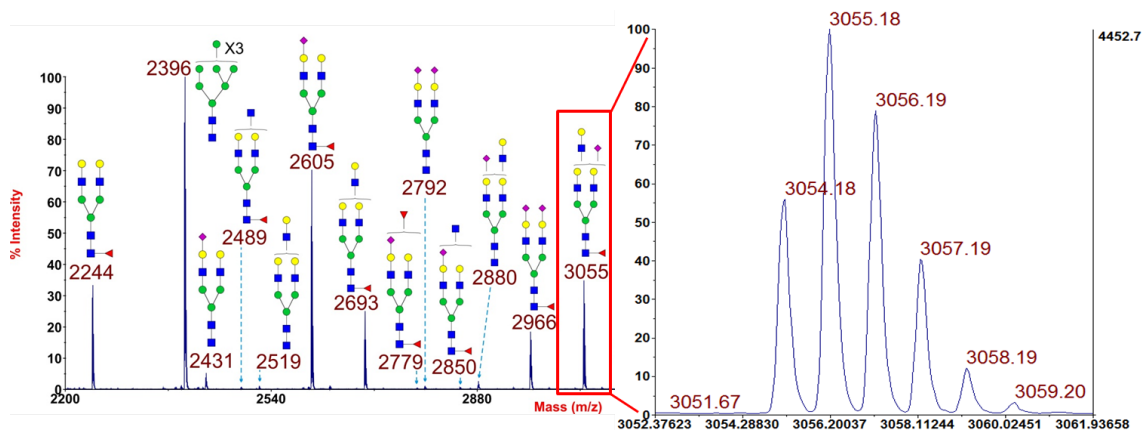
Figure 3.11 Annotated MALDI-TOF/TOF MS/MS spectrum of permethylated N-glycan at m/z 3055 in NHDF cells (A), the relative intensity of the same glycan in the MALDI-TOF MS spectrum (B)

Data were acquired in the form of $[M+Na]^+$ ions. See legends to Figure 3.2 and Figure 3.5 for explanation of structure assignments. In Panel A, only peaks of interest were labelled, the biantennary (bi) and triantennary (tri) characteristic ions are labelled in red and black frame respectively, RI is short for relative intensity. In B, the isotope peak of the glycan at m/z 3055 has been magnified. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

3.3.1.7. Summary

Complex N-glycans in both HeLa and NHDF cells have been quantified using the above method; the result (Figure 3.12) showed that before sialidase S digestion, the proportion of the glycan terminated with Gal was approximately 54% in NHDF cells, which is higher than that in HeLa cells (approximately 30%). After sialidase S treatment, the percentage of the glycan terminated with α2,6 linked NeuAc in HeLa cells is around 17%, which is significantly higher than the 3% shown in NHDF cells; sialidase A digestion removed all NeuAc residues supporting the glycomic assignments.
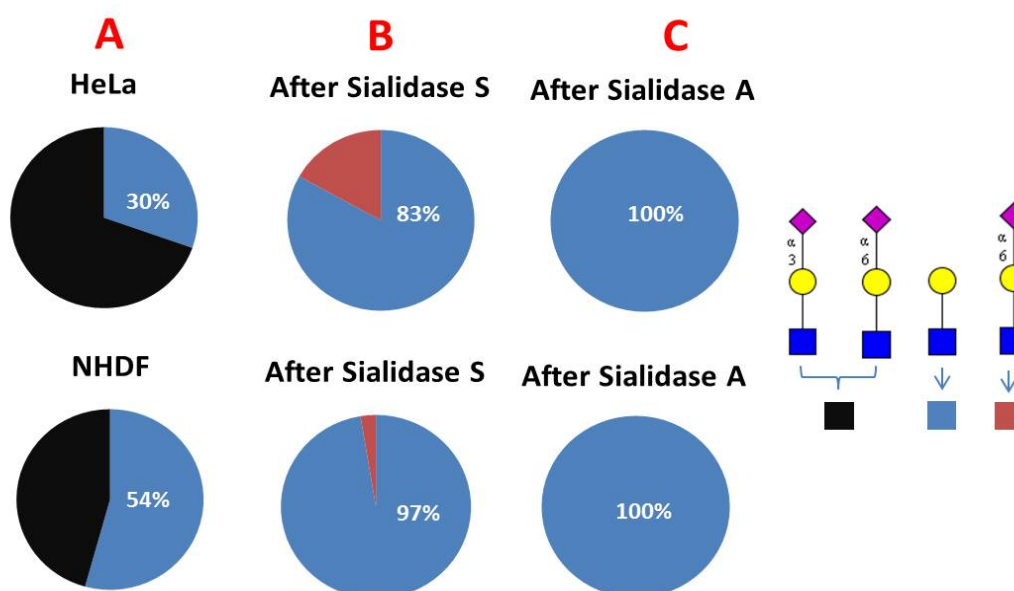
Figure 3.12 A comparison of the relative intensities of LacNAc antenna and sialylated LacNAc antenna in all complex glycans in HeLa and NHDF (A), sialidase S digested (B) and sialidase A digested (C)

Black colour stands for relative intensity of α2,3 and α2,6 sialylated LacNAc antenna; blue colour stands for relative intensity of LacNAc antenna; red colour stands for relative intensity of α2,6 sialylated LacNAc antenna. ■ GlcNAc, ● Gal, ◆ NeuAc.

3.3.2. HeLa and NHDF O-glycans

3.3.2.1.   MALDI-TOF MS analysis of the O-glycans of HeLa cells and NHDF

Duplicate analyses of O-glycans were also carried out to check the reproducibility and whether combining duplicates for the more detailed structural research would be valid. High quality MALDI data were obtained for the O-glycans from HeLa and NHDF cells. Representative MALDI-TOF MS spectra of HeLa and NHDF cell O-glycans are shown in Figure 3.13. The profile showed that in HeLa cells the glycan structures are core 1 (Galβ1,3GalNAc-O) type; monosialylated and disialylated core 1 were observed at m/z 895 and 1257 respectively. While in NHDF cells the glycan structures are core 1 (Galβ1,3GalNAc) and core 2 (Galβ1,3(GlcNAcβ1,6)GalNAc) types: in addition to observing monosialylated and disialylated core 1 at m/z at 895 and 1257 respectively, three core 2 derived glycans m/z 984, 1345 (monosialylated) and 1706 (disialylated) were also detected. In NHDF cells, core 1 derived glycans account for approximately 79.5% of all the O-glycans.
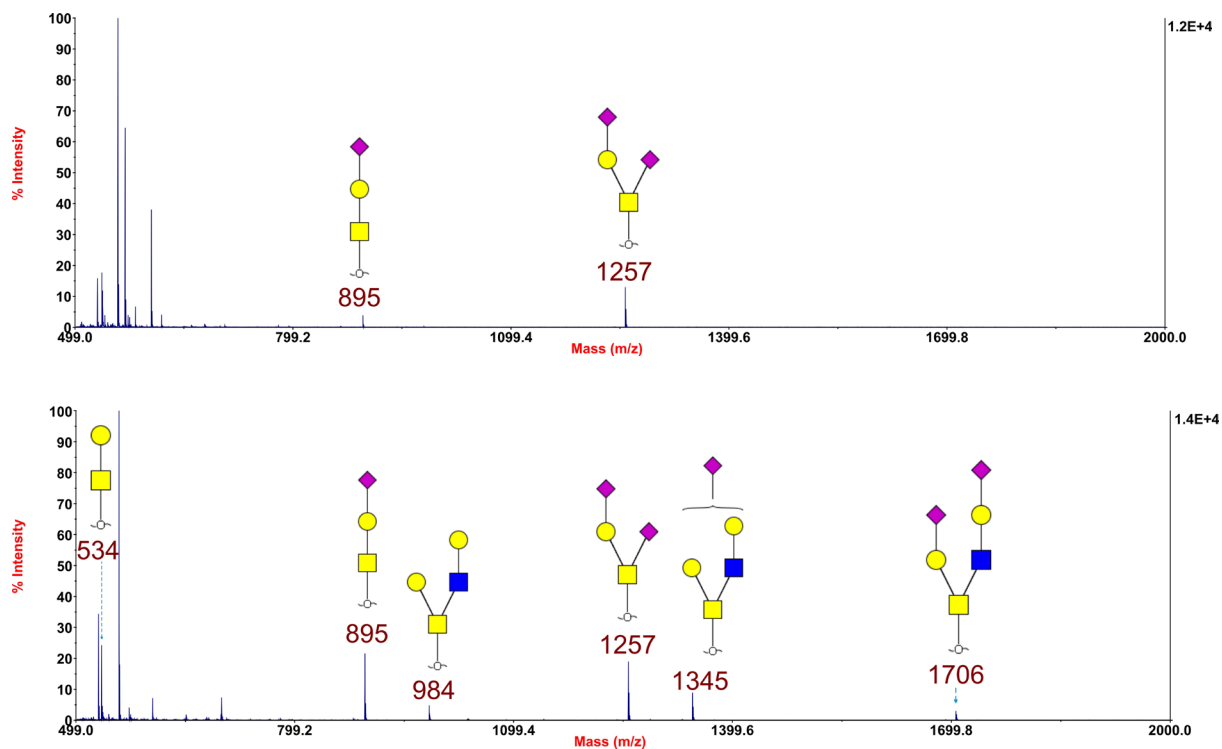
Figure 3.13 Annotated MALDI-TOF MS spectra of permethylated O-glycans in Hela (top) and in NHDF cells (bottom)

Profiles were obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]⁺. Putative structures are based on the molecular weight, O-glycan biosynthesis pathway and MS/MS data. ■ GalNAc, ■ GlcNAc, ○ Gal, ◆ NeuAc.

According to the O-glycan biosynthesis pathway, all O-glycan structures displayed in Figure 3.13 are unambiguous except the structures of the glycans at m/z 895 and 1345. The NeuAc in the glycan at m/z 895 can be either linked to the Gal or to the GalNAc. The NeuAc in the glycan at m/z 1345 glycans can be either linked to the Gal or to the Gal-GlcNAc. The position of sialylation was established by MS/MS where there were ambiguities.

As shown in Figure 3.14, in both HeLa and NHDF cells the Gal was confirmed to be 3-linked to the GalNAc as the signal at m/z 620 corresponds to an eliminated ion O-Gal-NeuAc which was derived from the elimination occurring at the C3 position of GalNAc. The presence of the fragment ion at m/z 620 and the absence of the fragment ions at m/z 259 (losing a single Gal via elimination) and 659 (the product of loss of a single Gal via elimination) showed that the NeuAc in the glycan at m/z 895 is linked to the Gal.
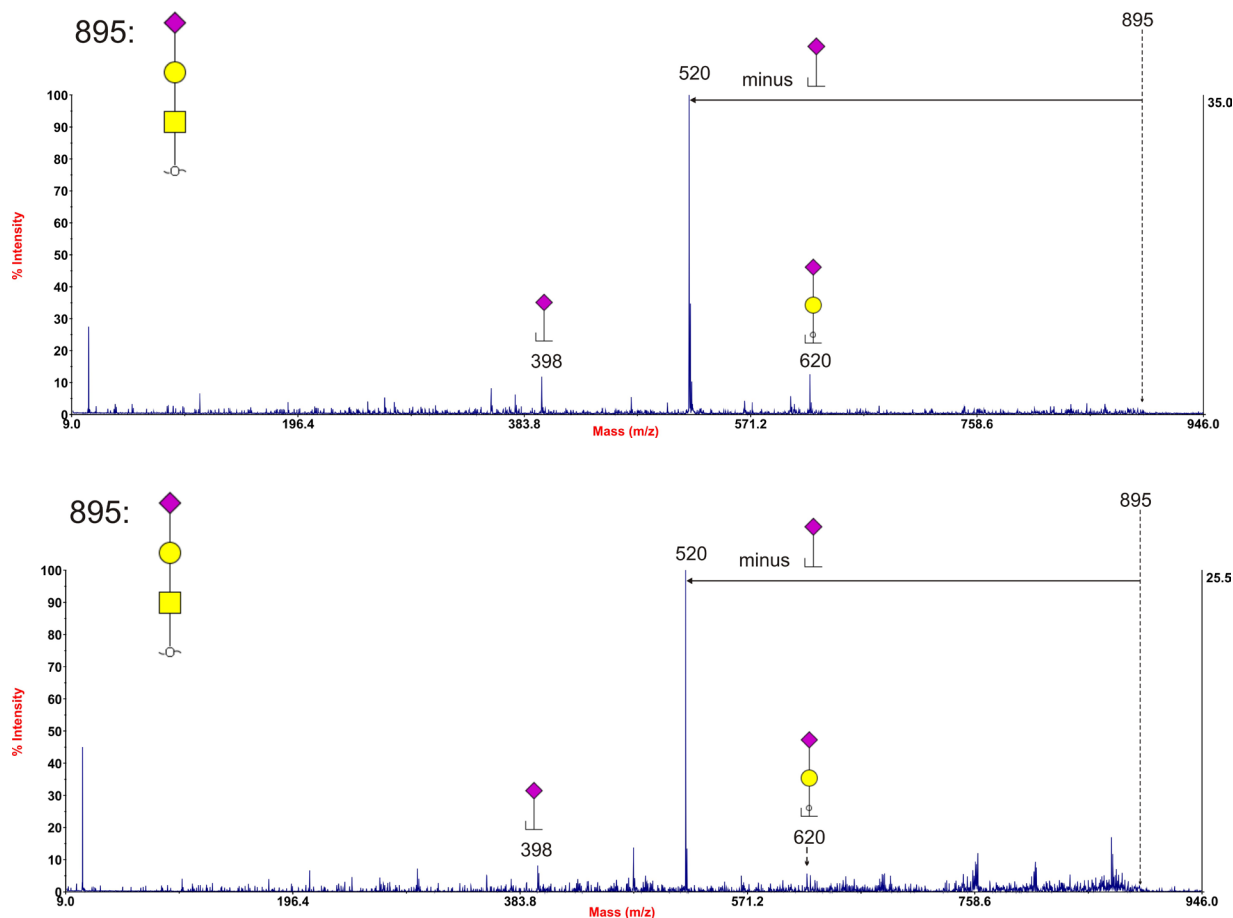
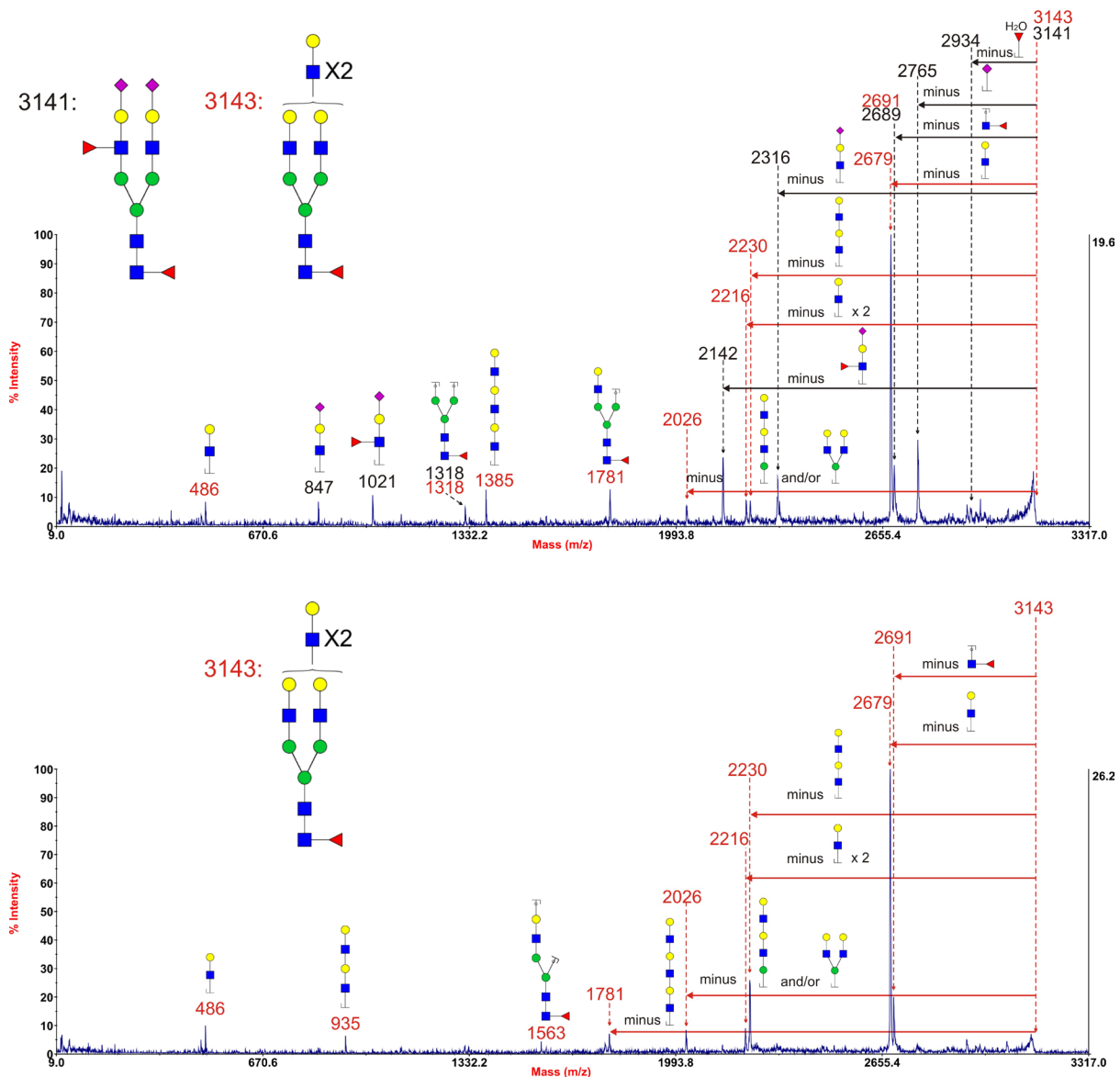**Figure 3.14 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated O-glycan at m/z 895 in HeLa cells (top) and NHDF cells (bottom)**

Data were acquired in the form of [M+Na]⁺ ions. See legend to Figure 3.5 for explanation of structure assignments. ▢ GalNAc, ◯ Gal, ◆ NeuAc.

As shown in Figure 3.15, in NHDF cells the Gal was confirmed to be 3-linked to the GalNAc as the signal at m/z 620 corresponds to an eliminated ion O-Gal-NeuAc which was derived from the elimination occurring at the C3 position of GalNAc, its concurrent ion at m/z 747 was also observed. In addition, the fragment ion at m/z 881 corresponds to a loss of LacNac, this demonstrates the presence of LacNAc. All these implied that the NeuAc was linked to the Gal. However, the molecular ion at m/z 520 corresponds to a loss of sialylated LacNac, this showed that the NeuAc was linked to the Gal-GlcNAc. Taken together, the MS/MS data provided evidence that the NeuAc was linked to either Gal or Gal-GlcNAc, which is indicated by drawing this NeuAc outside the bracket.
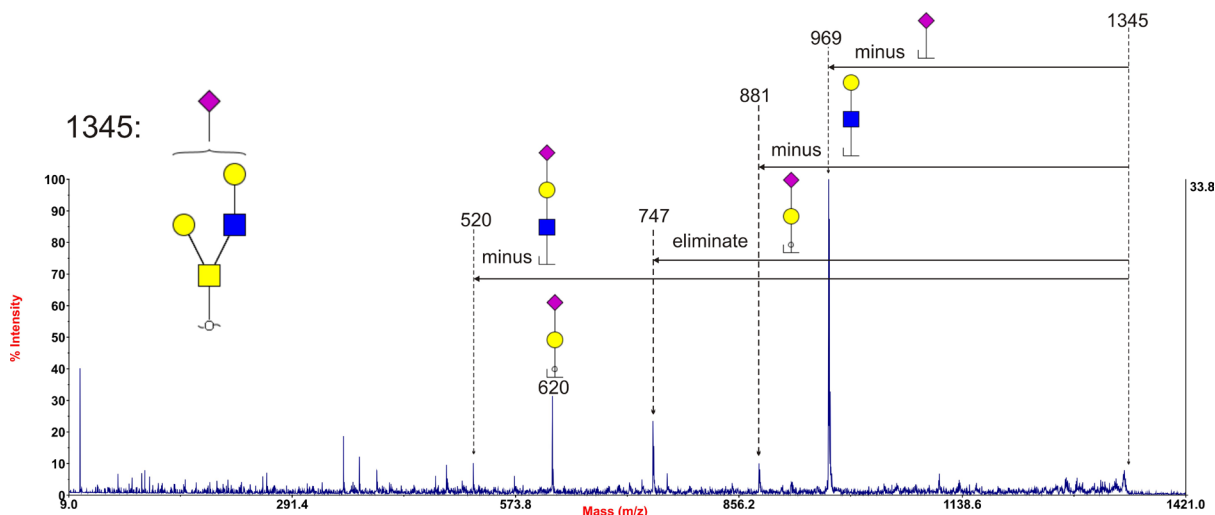
Figure 3.15 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated O-glycan at m/z 1345 in NHDF cells

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of structure assignments. ☐ GalNAc, ■ GlcNAc, ○ Gal, ◆ NeuAc.

3.3.2.2. *Sialidase S digestion of the O-glycans of HeLa and NHDF cells*

Sialidase S digestion was carried out on O-glycans to determine the NeuAc linkage. Digestion of both HeLa and NHDF cells with sialidase S resulted in partial desialylation (Figure 3.16). The core 1 structure (m/z 534) and monosialylated core 1 structure (m/z 895) shown in HeLa cells (Figure 3.16 top) were originally from the glycans at m/z 895 and 1257 in Figure 3.13 (top) respectively. A similar phenomenon was also observed in Figure 3.13 (bottom) and Figure 3.16 (bottom). In addition, the core 2 structure (m/z 984) shown in NHDF cells (Figure 3.16 bottom) was derived from the glycans at m/z 984, 1345 and 1706 in Figure 3.13 (bottom).

This indicated that both HeLa and NHDF contain α2,3 linked NeuAc. The following figure showed that a small amount of core 2 containing glycan (m/z 984) was observed in HeLa cells after sialidase S digestion, but this type of glycan was not detected prior to sialidase S treatment due to its low abundance.
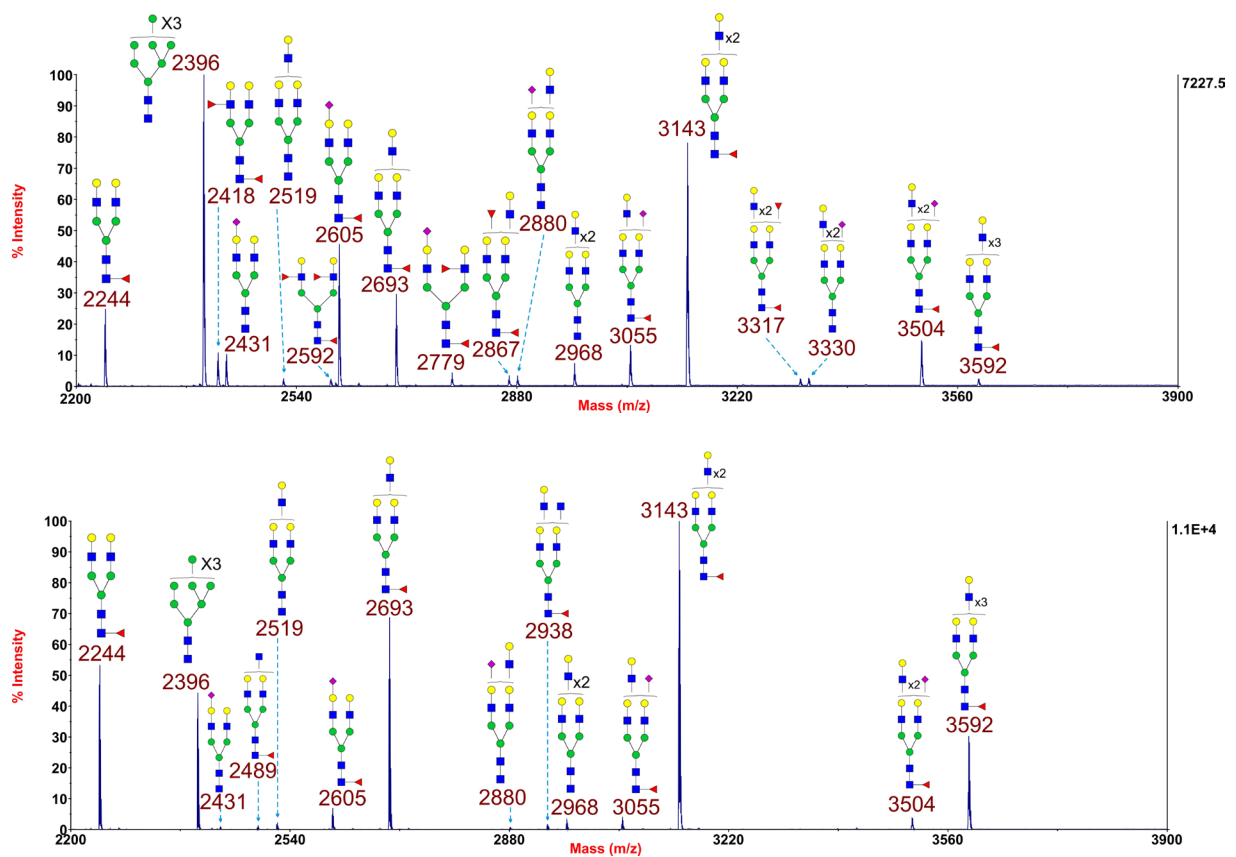
122
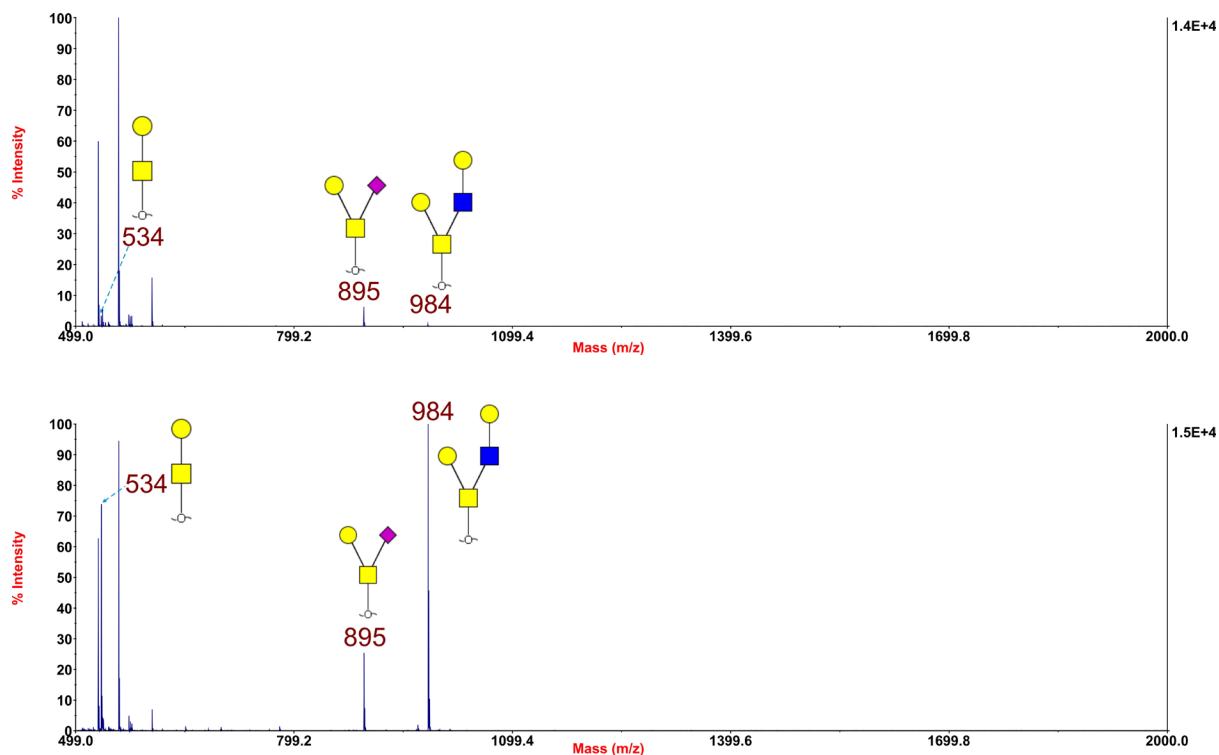
Figure 3.16 Annotated MALDI-TOF MS spectra of permethylated sialidase S treated O-glycans in Hela cells (top) and NHDF (bottom)
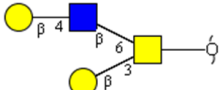
Profiles were obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.13 for explanation of structure assignments. ▢ GalNAc, ▪ GlcNAc, ◯ Gal, ◆ NeuAc.

3.3.2.3.   *Sialidase A digestion of the O-glycans of HeLa and NHDF cells*

Treatment with sialidase A led to a complete desialylation of the O-glycans in both cell types. Both core 1 and core 2 type glycans were observed after the desialylation (see Table 3.2).

Table 3.2 O-glycan structures observed in the MALDI-TOF MS spectra of Hela and NHDF after sialidase A digestion

Glycans were obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. Glycan structures and linkages are drawn based on the molecular weight, O-glycan biosynthesis pathway and MS/MS data. The rough percentage of the glycan relative abundance to the whole group is also indicated. ☐ GalNAc, ■ GlcNAc, ○ Gal.

m/z	Structure	After sialidase A digestion	
		HeLa cells (percentage)	NHDF cells (percentage)
534		96.0%	60.4%
984		4.0%	39.6%

3.3.2.4.   *Summary*

In both HeLa and NHDF cells the O-glycans are either terminated with NeuAc or with uncapped Gal. Comparison of the relative abundances of the O-glycans (Table 3.3) showed that in HeLa cells O-glycans were sialylated, the monosialylated core 1 glycan accounted for approximately 22.9% of all the O-glycans, the remaining was the disialylated core 1 glycan. A small amount of core 2 containing glycan was only observed after sialidase digestions (see sections 3.3.2.2 and 3.3.2.3). It also showed that in NHDF cells, the relative abundances of non- mono- and disialylated glycans accounted for 35.7%, 37.4% and 26.9% of all the O-glycans respectively. The relative abundance of disialylated glycans is greater in HeLa cells. After sialidase A digestion, the abundances of cores 1 and 2 (containing) glycans accounted for 60.4% and 39.6% respectively.

Table 3.3 O-glycan structures observed in the MALDI-TOF MS spectra of Hela and NHDF

All glycans are permethylated and $[M+Na]^+$. Glycan structures and linkages are drawn based on the molecular weight, O-glycan biosynthesis pathway and MS/MS data. ND, not detected. The ratio of the glycan relative abundance to the whole group is indicated using *; *=minor (<20%), **=medium (20-50%), ***=major (>50%). ■ GalNAc ■ GlcNAc, ● Gal, ◆ NeuAc.

m/z	Structures	Ratio in					
		HeLa	NHDF	HeLa, Sialidase S treated	NHDF, Sialidase S treated	HeLa, Sialidase A treated	NHDF, Sialidase A treated
534		ND	**	**	**	***	***
895		**	**	ND	ND	ND	ND
895		ND	ND	***	*	ND	ND
984		ND	*	*	***	*	**
1257		***	**	ND	ND	ND	ND
1345		ND	*	ND	ND	ND	ND
1706		ND	*	ND	ND	ND	ND

3.3.3. HeLa and NHDF glycolipid glycans

3.3.3.1. MALDI-TOF MS analyses of the glycolipid glycans of HeLa cells and NHDF

The glycolipid glycans were obtained following the procedure described in section 2.2. High quality MALDI data were obtained for the glycans from HeLa and NHDF cells. MALDI-TOF MS spectra of HeLa and NHDF cell glycolipid glycans are shown in Figure 3.17 and Figure 3.18. The spectra show that nearly all glycans detected in both HeLa and NHDF are sialylated with NeuAc. In HeLa cells glycolipid glycans with m/z 855, 943, 1101, 1305 and 1666 were detected in the 35% acetonitrile fraction. Their relative abundances are approximately 6.5%, 1.4%, 67.9%, 22.3% and 1.9% respectively. In addition, glycolipid glycans with m/z 1101, 1305 and 1666 were observed in the 50% acetonitrile fraction

(relative abundances 2.4%, 3.2% and 94.4% respectively). In NHDF cells glycolipid glycans with m/z 855, 943, 1217 and 1305 were detected in the 35% acetonitrile fraction, their relative abundances accounted for approximately 66.8%, 0.3%, 29.9% and 3.1% respectively. Additionally, glycolipid glycans with m/z 855, 1217, 1305 and 1666 were observed in the 50% acetonitrile fraction, their relative abundances accounted for 27.8%, 69.4%, 1.7% and 1.1% respectively. The glycan at m/z 1101 terminated with HexNAc is only observed in HeLa cells (see Table 3.4 for a compact summary).
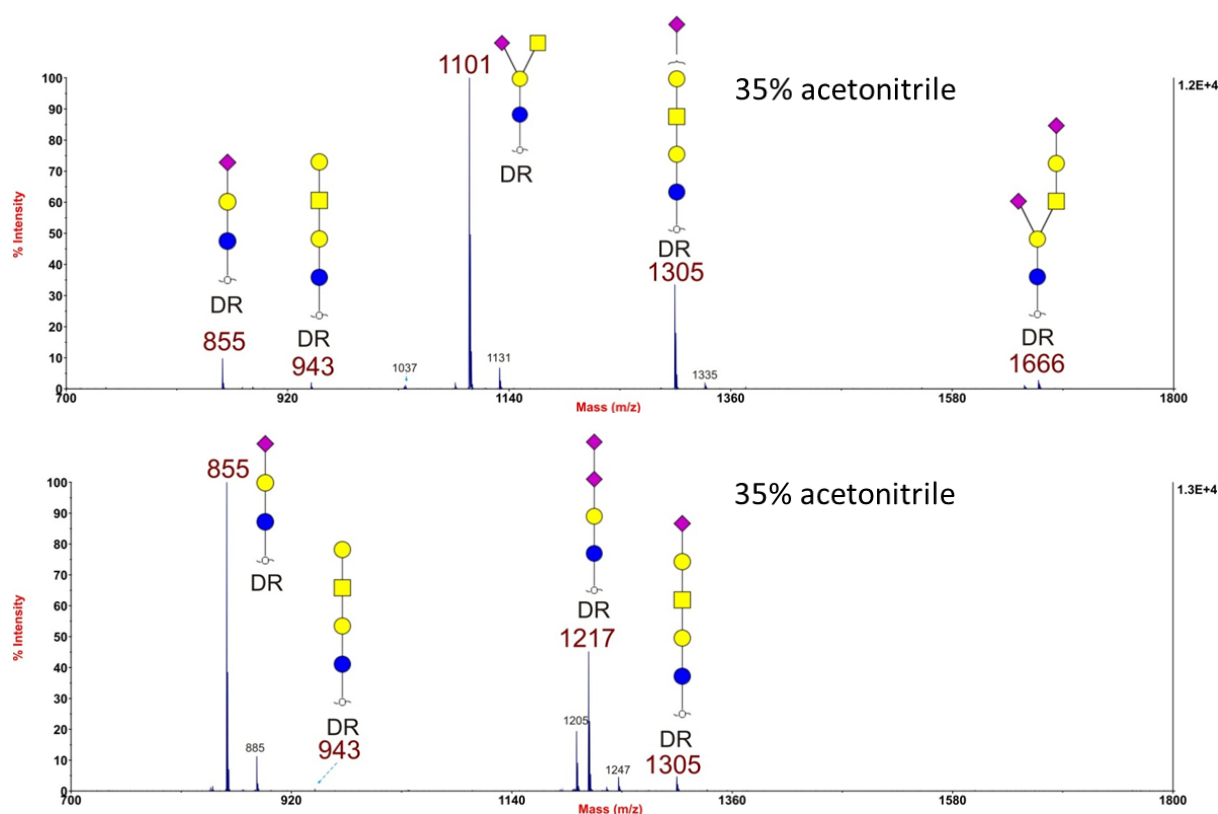
Figure 3.17 Annotated MALDI-TOF MS spectra of deuteroreduced, permethylated glycolipid derived glycans from Hela (top) and NHDF (bottom)

These profiles were obtained from the 35% acetonitrile fractions from a C18 Sep-Pak column. All glycans are deuteroreduced (DR) and in the form of [M+Na]$^{+}$. Putative structures are based on the molecular weight, glycolipid glycan biosynthesis pathway and MS/MS data. ☐ GalNAc, ● Glc, ○ Gal, ◆ NeuAc.
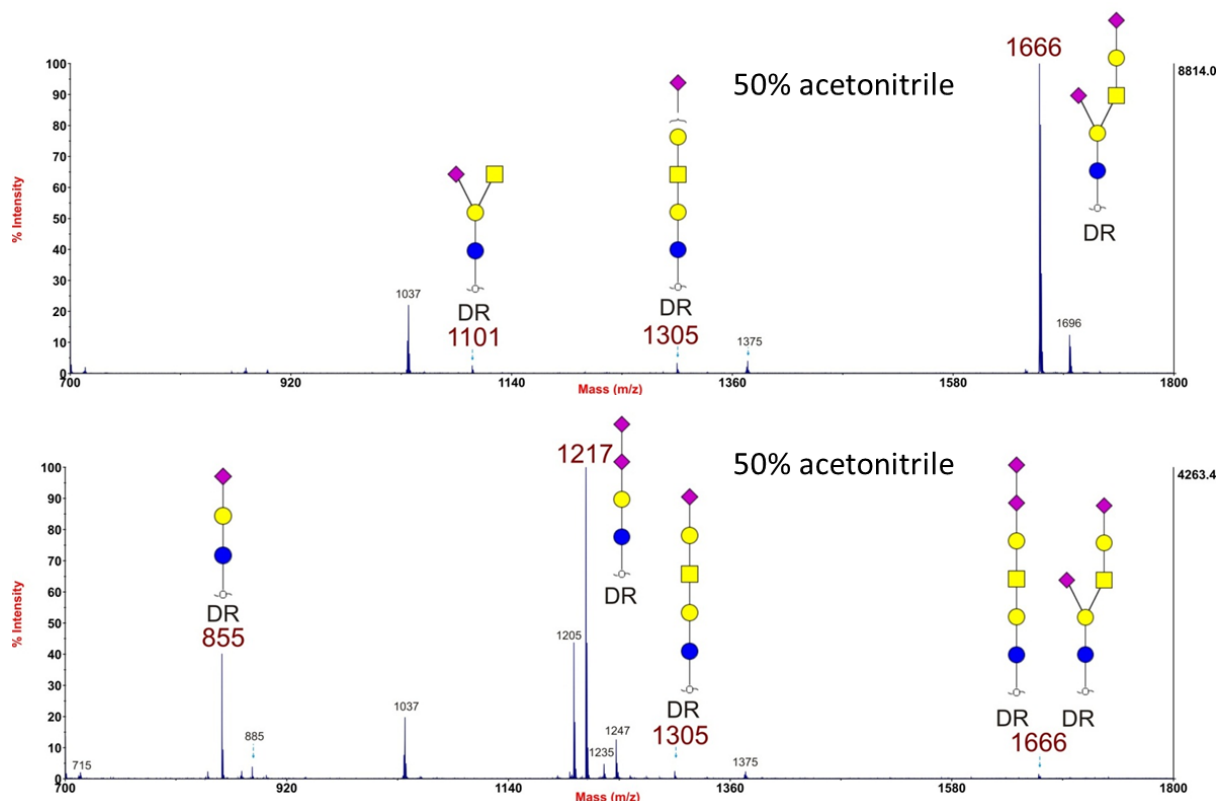
Figure 3.18 Annotated MALDI-TOF MS spectra of deuteroreduced, permethylated glycolipid derived glycans from Hela (top) and NHDF (bottom)

These profiles were obtained from the 50% acetonitrile fractions from a C18 Sep-Pak column. All glycans are deuteroreduced (DR) and in the form of $[M+Na]^+$. See legend to Figure 3.17 for explanation of structure assignments. ☐ GalNAc, ● Glc, ○ Gal, ◆ NeuAc.

Peaks of interest were chosen for further characterization via MALDI-TOF/TOF MS/MS analysis. For instance, the glycan at m/z 1305 (the position of the NeuAc could not be determined).

As shown in Figure 3.19 (top), the presence of the fragment ion at m/z 847 and its concurrent ion at m/z 480 showed that the NeuAc is linked to the terminal Gal. The presence of the fragment ion at m/z 486 and the presence of its concurrent ion at m/z 841 showed that the NeuAc is linked to the internal Gal. In addition, the presence of the double-cleavage ion at m/z 466 could also support this. All these showed that the glycan at m/z 1305 in HeLa cells possessed two structures; the NeuAc is either linked to the terminal Gal or to the internal Gal. In contrast to the HeLa data, MS/MS of m/z 1305 from NHDF cells showed fragment ions that correspond solely to one structure (Figure 3.19 bottom), in which the NeuAc is linked to terminal Gal.

127
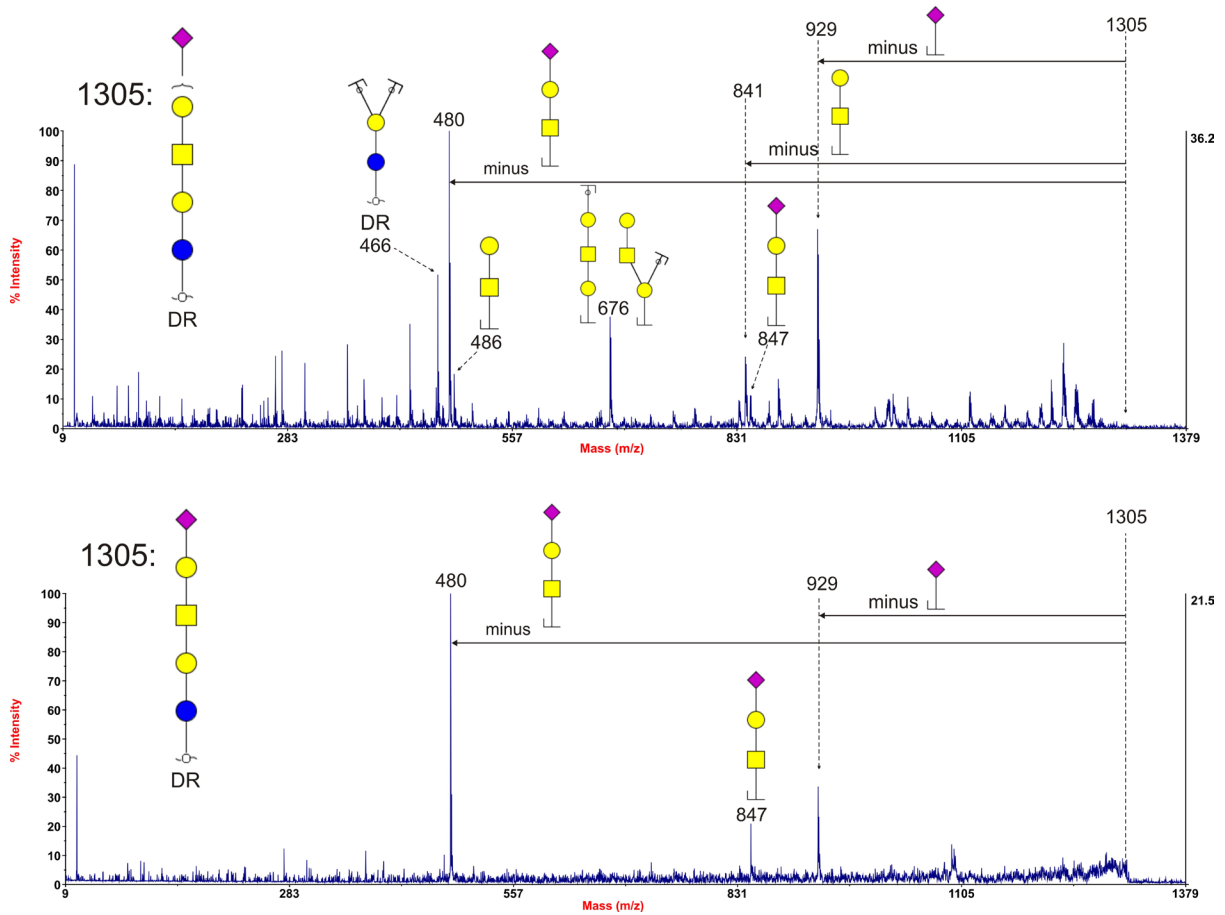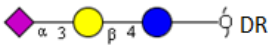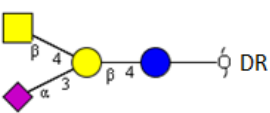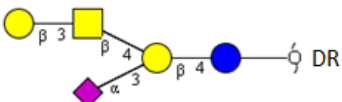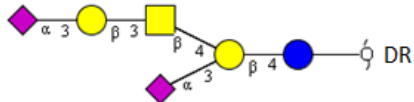
Figure 3.19 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated glycolipid glycan at m/z 1305 in HeLa cells (top) and NHDF cells (bottom)

Data were acquired in the form of [M+Na]$^+$ ions. DR, deuteroreduced. See legend to Figure 3.5 for explanation of structure assignments. ▢ GalNAc, ● Glc, ○ Gal, ◆ NeuAc.

The following table summarises the structural conclusions which are extracted from the MS and MS/MS analyses, taking into account glycolipid glycan biosynthetic considerations.

Table 3.4 Structures of glycans derived from glycolipids observed in the MALDI-TOF MS spectra of Hela and NHDF cells

All glycans are deuteroreduced (DR), permethylated and in the form of $[M+Na]^+$. Glycan structures are drawn based on the molecular weight, glycolipid glycan biosynthesis pathway and MS/MS data; ND, not detected. The ratio of the glycan relative abundance to the whole group is indicated using *. *=minor (<20%), **=medium (20-50%), ***=major (>50%). □ GalNAc, ● Glc, ○ Gal, ◆ NeuAc.

m/z	Structures	Ratio in 35% acetonitrile		Ratio in 50% acetonitrile	
		HeLa	NHDF	HeLa	NHDF
855	GM3	*	***	ND	**
943	GA1	*	*	ND	ND
1101	GM2	***	ND	*	ND
1217	GD3	ND	**	ND	***
1305	GM1a	**	ND	*	ND
1305	GM1b		*		*
1666	GD1a	*	ND	***	*
1666	GD1c	ND		ND	

3.3.3.2. Sialidase digestion of the glycolipid glycans of HeLa cells

Sialidase digestions were only carried out on the glycolipid glycans from HeLa cells due to the fact that the amount of NHDF cells was limited. A new batch of HeLa cells was used for this enzymatic digestion. High quality MALDI data were obtained. As shown in Figure 3.20, digestion of the glycans with sialidase S resulted in partial desialylation. Thus the sialylated glycan at m/z 1101 was still observed while the glycan at m/z 1666 was absent. One structure giving the original signal at m/z 1305 was absent, whilst the other was still observed. A new glycan, the m/z of which is 943, was detected. All these confirmed the α2,3 linkage of the

peripheral NeuAc in glycans at m/z 1305 and 1666. Sialidase A digestion resulted in a complete loss of NeuAc supporting the glycomic assignments.
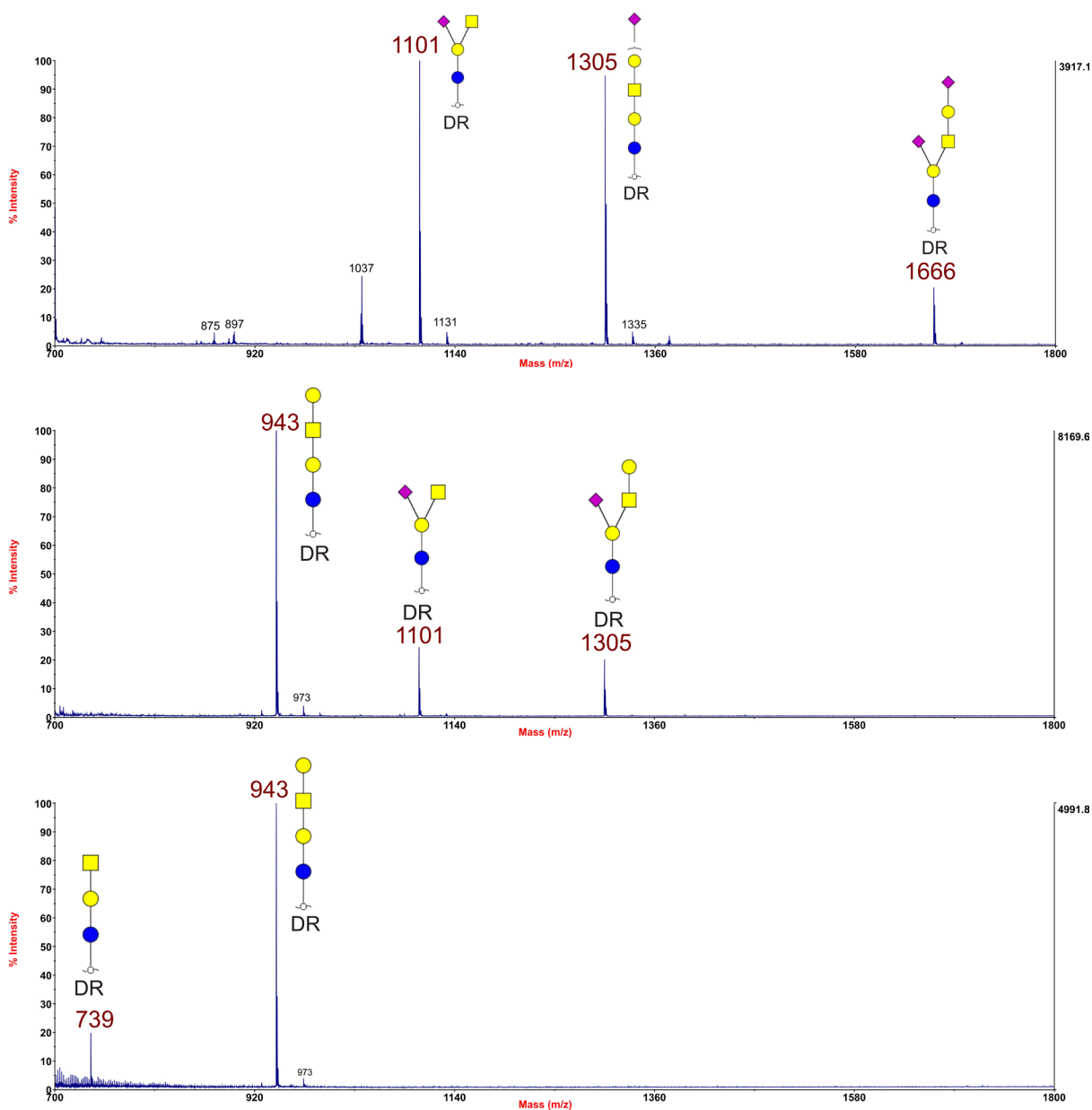
Figure 3.20 Annotated MALDI-TOF MS spectra of deuteroreduced, permethylated glycolipid derived glycans from Hela cells, untreated (top), sialidase S treated (middle), sialidase A treated (bottom)

These profiles were obtained from the 35% acetonitrile fractions from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.17 for explanation of structure assignments. ☐ GalNAc, ● Glc, ○ Gal, ◆ NeuAc.

The result of glycomic analyses showed similar sialylation in the glycolipid glycans in the two cell types.

3.4. Discussion

Although Van Damme and Shang had investigated these SNAs using glycan array and HeLa and NHDF cells, they did not determine the cell glycomes; this is essential as it could bridge what they did determine (glycan array result, SNA cytotoxicity result (Table 3.5), etc.) (Shang et al., 2015; Shang and Van Damme, 2014) and form a more complete story on these SNAs. Therefore, the glycomic analysis was carried out.

Table 3.5 Comparison of LC50 values for the *S. nigra* proteins in HeLa and NHDF cell lines

Data are shown as means ± SE based on 4 replications per treatment, and each experiment was repeated 3 times. LC50, lethal concentration that kills half of the sample population. The concentration of SNAs used for the cell viability test is in the range of 0.1-2 μM. The concentrations (>2 μM) were obtained via calculated according to the trend of the dose response curve. Since these predict LC50 concentrations (>2 μM) were not tested, accurate values were not given. Therefore, this conclusion, " the LC50 value is > 2 μM ", was made (Shang et al., 2015).

LC50 (μM)					
Time/Cell line	SNA-I	SNA-II	SNA-IV	SNA-V	SNLRP
24h Hela	1.57 ± 0.319	0.72 ± 0.106	>2	0.74 ± 0.083	>2
48h Hela	0.43 ± 0.041	0.27 ± 0.03	1.932 ± 0.202	0.11 ± 0.014	1.85 ± 0.437
24h NHDF	>2	>2	>2	>2	>2
48h NHDF	1.70 ± 0.35	2.00 ± 0.77	>2	0.74 ± 0.118	>2

The glycomic analyses of HeLa cells and NHDF cells demonstrated a clear correlation with the results from cytotoxicity and internalization studies, which indicates the importance of protein-carbohydrate interactions for the proteins entering the cells. Particularly, the result of N- and O-glycomic analyses showed that compared to NHDF cells the overall sialylation level of N- and O-glycans in HeLa cells is greater (see Figure 3.12 and Figure 3.13), which could explain why SNA-IV was more toxic towards HeLa cells. Similarly, significantly more α2,6 linked NeuAc residues in the glycoprotein glycans from HeLa cells (see Figure 3.12) could explain the greater cytotoxicity of SNA-I to HeLa cells. In addition, SNA-II and SNA-V also showed more cytotoxicity to the HeLa cells (see Table 3.5), which was possibly due to the fact that these two SNAs preferred binding to O-glycan core 1 glycan (see Table 1.7).

Moreover, compared to other SNAs, SNA-V is the only protein that revealed remarkable cytotoxicity to NHDF cells (see Table 3.5), and this could be caused by the higher abundance of terminal Gal residues in these cells. As shown in Table 1.7, SNLRP prefers binding to the structure GlcNAcβ1,4GlcNAc which is the chitobiose core moiety (Barry et al., 2013). Therefore, this SNA should interact with all N-glycans as this disaccharide is a core structure for all N-glycans. However, the LC50 values in the cytotoxicity experiment for SNLRP are very great (see Table 3.5), implicating low binding to the cell surface. More importantly, this was consistent with the result of the SNA internalization experiment performed using HeLa cells. In the experiment the amount of SNLRP internalized by HeLa cells was extremely low (Shang et al., 2015), and this suggested that the binding of SNLRP to the chitobiose core moiety is probably blocked as the structure GlcNAcβ1,4GlcNAc is cryptic due to capping structures.

Although the MS data provided glycan profiles of HeLa and NHDF cells and new insights on potential carbohydrate binding sites which associate with cell binding/uptake for SNAs, it is evident that this interaction is complicated and not adequate to explicate the cytotoxicity of the proteins under study. For instance, although the amount of terminal Gal on HeLa cells is obviously lower than that of sialic acid, the reason why the SNA-V and SNA-II are more toxic to HeLa cells than SNA-I is still unknown. A possible explanation for this could be differential intracellular trafficking of the proteins. It has been reported that the cytotoxicity of some type 2 RIPs, such as ricin and abrin, relies on the binding of the B chain to glycoconjugates at the cell surface, as such promoting cellular uptake of the RIPs (Sandvig et al., 1978; Lannoo and Van Damme, 2014; Shang et al., 2015). Incubation of mammalian cells with ricin leads to it being endocytosed and then transported to the trans-Golgi network, followed by retrogradely transported from the Golgi to the ER in which the disulphide bond is cleaved, after that the A chain gets into the cytoplasm where it finally refolds and inhibits protein synthesis via removing a key adenine from the 28S ribosomal subunit (Sandvig et al., 1991; Lord and Roberts, 1998; Spooner and Lord, 2015). However, the intracellular trafficking of SNA V is different; it enters the cytosol without passing the Golgi and ER (Barbieri et al., 2004; Battelli et al., 1997; Shang et al., 2015). Therefore, it is possible that these SNAs follow different intracellular trafficking and exhibit the cytotoxicity via different ways.

It has been reported that RIPs are able to induce cell death via apoptosis (Das et al., 2012). Since SNA-I, SNA-V and SNLRP are type-2 RIPs, they should be able to cause cell

apoptosis. In addition, it has been reported that SNA-II can cause apoptosis (Shahidi-Noghabi et al., 2010), and both SNA-II and SNA-IV possess one lectin chain, this implies that SNA-IV is probably able to cause apoptosis as well. Therefore, all these SNAs are able to cause apoptosis via which HeLa and NHDF cells could be killed in the experiment. However, it is not clear that how SNAs trigger the apoptosis. It is reported that SNA-I and SNA-II could trigger caspase-3 like activities in cells (Shahidi-Noghabi et al., 2010), and caspase-3 is a cell death protease and it plays an essential role in apoptosis (Porter and Janicke, 1999). It is therefore possible that the binding of SNAs to cells yields a stimulus which subsequently activates the caspase pathway resulting in apoptosis.

Sambucus species have been used for medicinal purposes since Pedanius Dioscorides's (a Greek physician, pharmacologist and botanist) period (40-90 A.D.). Among these, *Sambucus nigra* lectins are the most relevant species researched (Tejero et al., 2015). However, its application has been restricted to some extent due to the fact that the proteins are bioactive and the ingestion of these compounds could result in deleterious effects. In the last few years, the chemical and pharmacological characteristics of *Sambucus* species have been investigated. SNA-V (Nigrin b) can be used for the construction of conjugates which target the transferrin receptor, and this receptor is over-expressed in cancer cells (Citores et al., 2002). Although the biological role of proteins derived from *Sambucus* remains unclear, they are speculated to be involved in helping plants against insects and viruses (Tejero et al., 2015). Therefore, these SNAs could be potentially used in agriculture to protect crops against plant insects and viruses.

This study reports for the first time the glycolipid glycans in NHDF cells by MALDI-TOF MS and MALDI-TOF/TOF MS/MS. Although a paper reported N- and O-glycan profile of NHDF in 2012 (Engelstaedter et al., 2012), half of their sample amount was used and 31 more complex N-glycans were detected in this glycomic analysis. Therefore the analysis here is more sensitive and complete, In addition, the detection of the potential sialyl-Lewis X in the N-glycans from HeLa cells supports the accuracy of the glycomic analyses as sialyl-Lewis X is expressed in cervical cancer (Engelstaedter et al., 2012). More work is required for uncovering the veil of the biological role and the mechanism of SNA proteins. For instance, plants can be genetically modified (being transfected with the gene that encodes SNA-I or SNA-II) to express the SNA. Theoretically, these transformants should possess enhanced insect resistance as these two SNAs can cause insect midgut cell apoptosis

(Shahidi-Noghabi et al., 2010). If this is working, it can be applied to tobacco plants, which will bring great profits.

Chapter 4

Mass spectrometric investigation of glycosylation in patients with muscular diseases

4. Mass spectrometric investigation of glycosylation in patients with muscular diseases

4.1. Introduction to the project

This project is in collaboration with Professor Hanns Lochmuller and Dr. Juliane Mueller at Newcastle University.

To test the hypothesis which is mutations in *GFPT1* may cause N-glycan branching variations and thus have an influence on protein glycosylation (see section 1.7.1), glycomic analysis was performed to determine whether protein glycosylation is generally impaired or modified in CMS patients with *GFPT1* mutations. The N-glycomes in myoblasts from healthy controls and patients with *GFPT1* mutations, as well as myotubes obtained by *in vitro* differentiation were rigorously characterised. Simultaneously, myoblasts and myotubes from other patients were also investigated in the same way to check whether this possible impairment or modification is unique in the *GFPT1* patients. Specifically, MS analyses could give general structural information of the glycans from healthy controls and patients. Since the focus is on glycan branching variations, potential branched glycans from the controls and patients were further analysed with MS/MS as this technique can determine glycan branch structures. Sialidase S digestion was only carried out in the myoblast and myotube samples from the *DOK7* patient to determine the sialic acid linkage due to the fact that the cell pellets (myoblast and myotube) from this patient were visually greater than others.

This investigation would provide a better understanding of glycosylation in neuromuscular junction disorders with the long term objective of establishing structure/function relationships. Additionally, potential glyco-biomarkers of congenital myasthenic syndrome (CMS) may be discovered when comparing glycan profiles of healthy controls, *GFPT1* patients and other muscular disease patients.

4.2. Sample details and sample processing

The detail of three batches of samples is shown in the following 3 Tables. Cell culture details are described in section 2.1.3.2.

Table 4.1 Detail of myoblasts and myotubes

"*GFPT1* patient" means a CMS patient with GFPT1 mutations; similarly for *DOK7* and *MTND5* patients. Each of the cell lines, except for *GFPT1* patient 3, has three pellets. The myotubes were differentiated from myoblasts. (The table has been changed as required: previous three tables have been combined)

First batch		
Myoblasts are from	**Cell number**	**Number of samples**
GFPT1 patient 1	Estimated to be 10 million	3
GFPT1 patient 2	Estimated to be 10 million	3
GFPT1 patient 3	Estimated to be 10 million	1
DOK7 patient	Estimated to be 10 million	3
MTND5 patient	Estimated to be 10 million	3
Second batch		
Myoblasts are from	**Cell number**	**Number of samples**
GFPT1 patient 1	Estimated to be 10 million	1
GFPT1 patient 2	Estimated to be 10 million	1
GFPT1 patient 3	Estimated to be 10 million	1
DOK 7 patient	Estimated to be 10 million	1
MTND5 patient	Estimated to be 10 million	1
Healthy control 1	Estimated to be 10 million	2
Healthy control 2	Estimated to be 10 million	2
LGMD2A patient	Estimated to be 10 million	2
Pompe disease patient	Estimated to be 10 million	2
Third batch		
Myotubes are from	**Cell number**	**Number of samples**
GFPT1 patient 1	unknown	1
GFPT1 patient 2	unknown	1
GFPT1 patient 3	unknown	1
Healthy control 1	unknown	1
Healthy control 2	unknown	1
DOK 7 patient	unknown	1
MTND5 patient	unknown	1
LGMD2A patient	unknown	1
Pompe disease patient	unknown	1

All samples were analysed using glycomic methodologies which have been described in section 2.2.

4.3. Results

4.3.1. Determination of optimal conditions for cell culture

Because sufficient sample for glycomic analysis could not be provided by human muscle biopsies, suitable cell culture conditions for producing cells are required to be established whilst the amount of FCS used in the culture medium should be minimised. The latter was important as it is known that glycans derived from FCS are frequently co-purified with cell-derived glycans during glycomic analyses (Monk et al., 2006).

4.3.1.1. *MALDI-TOF MS analysis of the N-glycans of the myoblasts cultured in the medium containing 5%, 10% and 15% serum*

Myoblast samples listed in Table 4.1 were analysed using glycomic methodologies. These myoblasts were cultured in medium with different FCS concentrations. Good quality MALDI-MS spectra were only yielded from myoblasts cultured in 15% FCS. Representative data are shown in Figure 4.1. The spectra demonstrate the presence of high mannose (m/z 1580, 1784, 1988, 2192 and 2396) and complex glycans (bi-, tri- and tetraantennary, e.g. m/z 2966, 3777 and 4587) in the cell line. The relative abundance of all complex glycans in top, middle and bottom panels which were obtained from myoblasts cultured in the medium supplemented with 5%, 10% and 15% FCS respectively accounted for approximately 9.5%, 29.6% and 45.4% respectively of all detected glycans. Especially, Top and middle panels showed that the high mass glycans were not as abundant compared to that in bottom panel. More detailed interpretations of figures will be shown in the next section.
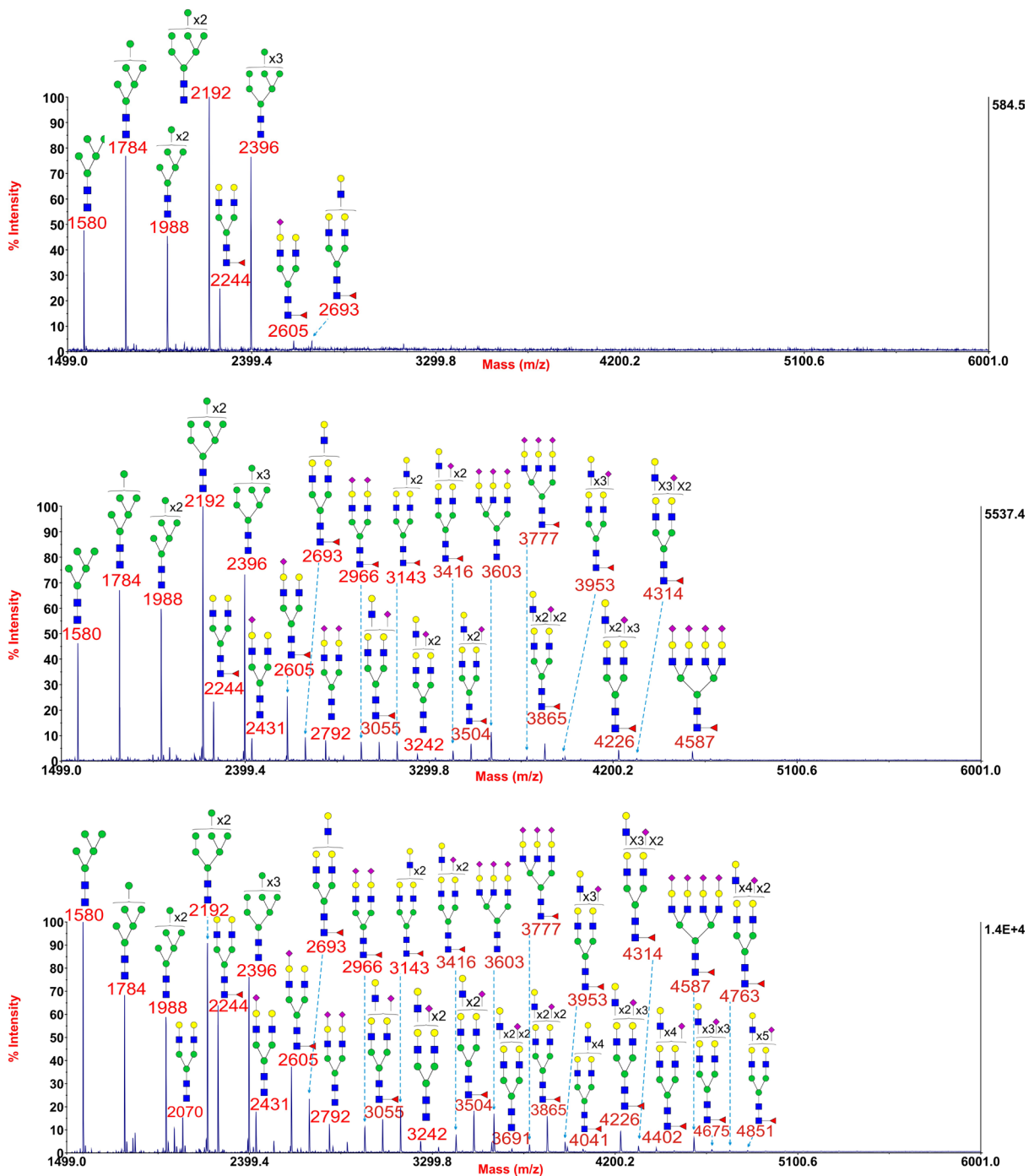
Figure 4.1 Annotated MALDI-TOF MS spectra of permethylated N-glycans from *DOK 7* patient myoblast cultured in the medium containing 5% (A), 10% (B) and 15% (C) FCS

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

According to the hypothesis (see section 1.7.1), the focus of this project is on N-glycan branching variation. Therefore, glycans with at least 2 LacNAc units are required to check the variation. Glycans with at least 2 LacNAc units were sufficiently abundant when cells grew in the medium containing 15% FCS, which enabled us to make firm conclusions about branching. In addition, one cell line (*GFPT1* patient 3) grew extremely slowly in the medium containing 5% and 10% FCS.Therefore, all cells that analysed in this project afterwards were cultured in the medium containing 15% FCS.

The glycans which lack core fucose shown in the spectra (Figure 4.1) are probably derived from FCS (Monk et al., 2006) and these glycans do not seem to be a problem as they are not dominating the spectra, thus the contamination will not have too much influence on the final result. In all the comparisons carried out afterward, the non-core fucosylated glycans are excluded.

4.3.2. Myoblast N-glycans

4.3.2.1. ***MALDI-TOF MS analysis of the N-glycans of the myoblasts from healthy controls and GFPT1 patients and other muscular disease patients***

Myoblast samples listed in Table 4.1 were analysed. High quality MALDI data were obtained for the N-glycans from all myoblast samples with the exception of *GFPT1* patient 3 whose myoblasts were difficult to culture. Combining with the five MALDI MS spectra (myoblasts cultured in 15% FCS containing medium) from the previous section a total of eighteen MALDI MS data sets were generated which included duplicate experiments. Representative MALDI-TOF MS spectra from a healthy control, a *GFPT1* patient and the *DOK7* patient are shown in Figure 4.2 top, middle and bottom panel, respectively.

The N-glycans comprised high mannose and complex glycans. High mannose glycan structures were the same as the ones mentioned in section 4.3.1.1. In addition to being bi-, tri- and tetraantennary (see section 4.3.1.1), the complex N-glycans in the myoblasts (Figure 4.2 and Figure 4.3) could be nonsialylated (e.g. m/z 2244, 2693 and 3143), monosialylated (e.g.

m/z 2605, 3055 and 3504), disialylated (e.g. m/z 2966, 3416 and 3865), trisialylated (e.g. m/z 3777 and 4226) and tetrasialylated (m/z 4587). No disease-unique glycan structure was identified. Some common characteristics of mammalian cell N-glycomes (Dell and Morris, 2001; Antonopoulos et al., 2012) were observed, such as core fucosylated GlcNAc, LacNAc antenna building blocks which in some cases are tandemly repeated to produce oligo-LacNAc extensions, and NeuAc capped antennae. The number of LacNAc units observed in the detected glycan was up to nine (m/z 5389, Figure 4.2 middle). Significantly, the N-glycan profile of healthy control 1 showed a broadly similar pattern to those of *GFPT1* patient 1 and *DOK7* patient.

Figure 4.2 Annotated MALDI-TOF MS spectra of permethylated N-glycans of myoblasts from healthy control 1 (top), *GFPT1* patient 1 (middle) and the *DOK7* patient (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are $[M+Na]^+$. The number indicated in the spectra is the mass to charge ratio (m/z) of the corresponding ion. Since the ion is monocharged, the value of m/z is equal to the molecular weight value of the glycan. Annotations are based on the molecular weight, N-glycan biosynthesis pathway and MS/MS data. Glycans at m/z 2966, 3777 and 4587 are clearly annotated, this is due to the fact that their structures are unequivocal because each antenna is capped with a sialic acid and thus they are homogeneous bi-, tri- and tetraantennary glycans. However, the glycan structure is not always as unequivocal as the glycan at m/z 2966 as biosynthetically non-fully sialylated glycan molecular ion species could be made up of mixtures of structural isoforms. Therefore, for those heterogeneous multiantennary structures with extended LacNAc repeats, the annotations are simplified throughout by using biantennary structures with the extensions and NeuAcs listed outside a bracket. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

Figure 4.3 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) of myoblasts from healthy control 1 (top), GFPT1 patient 1 (middle) and the DOK7 patient (bottom)

This figure amplifies the mass range (m/z 2900-4700) where the majority of tri- and tetra-antennary glycans are found in Figure 4.2. Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

MALDI data for the other patients and control are shown in Figure 4.4, Figure 4.5, Figure 4.6 and Figure 4.7; there was also no significant difference in the N-glycan profile among these samples.
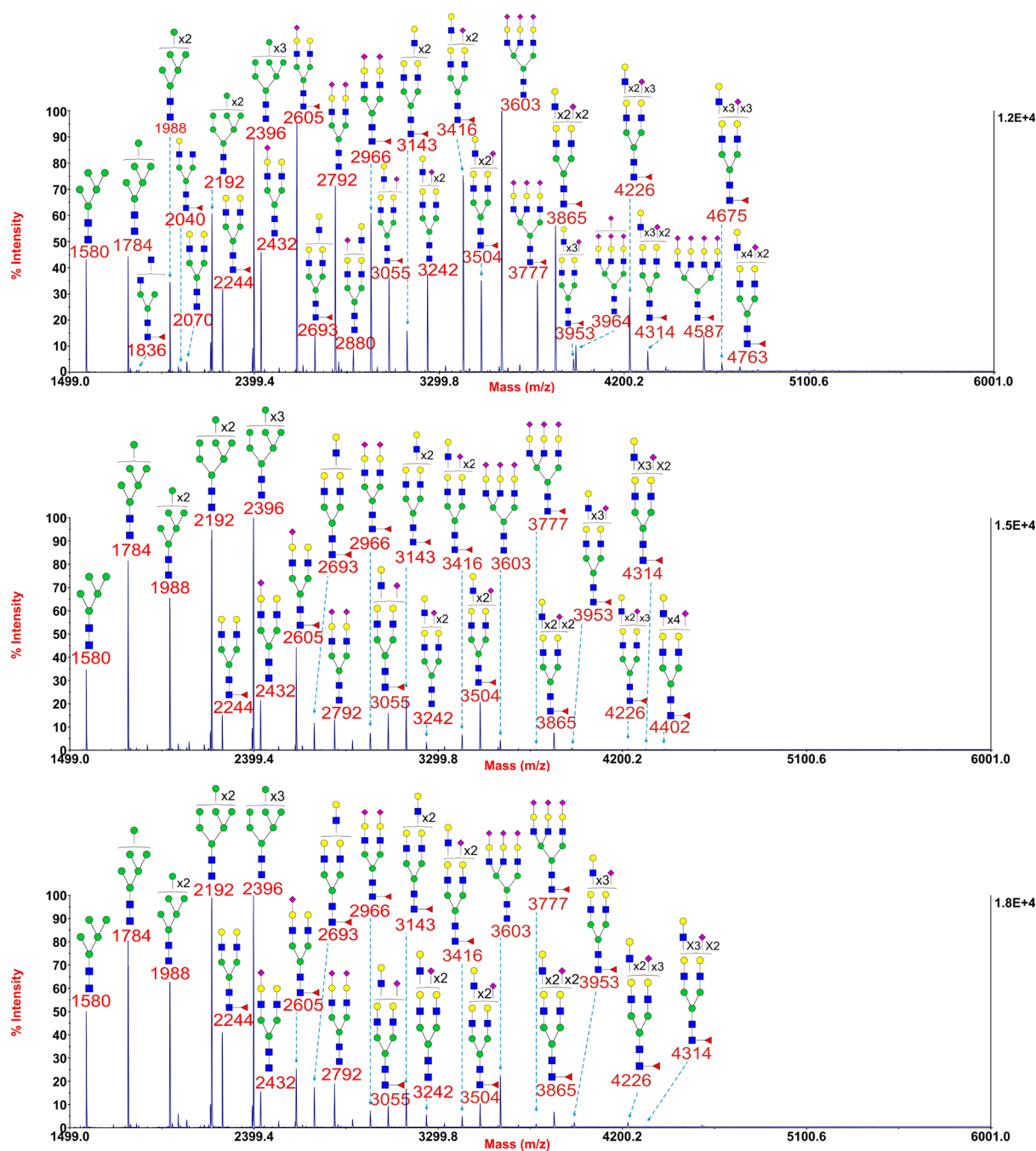
Figure 4.4 Annotated MALDI-TOF MS spectra of permethylated N-glycans from myoblasts of healthy control 2 (top), *GFPT1* patient 2 (middle), the *MTND5* patient (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
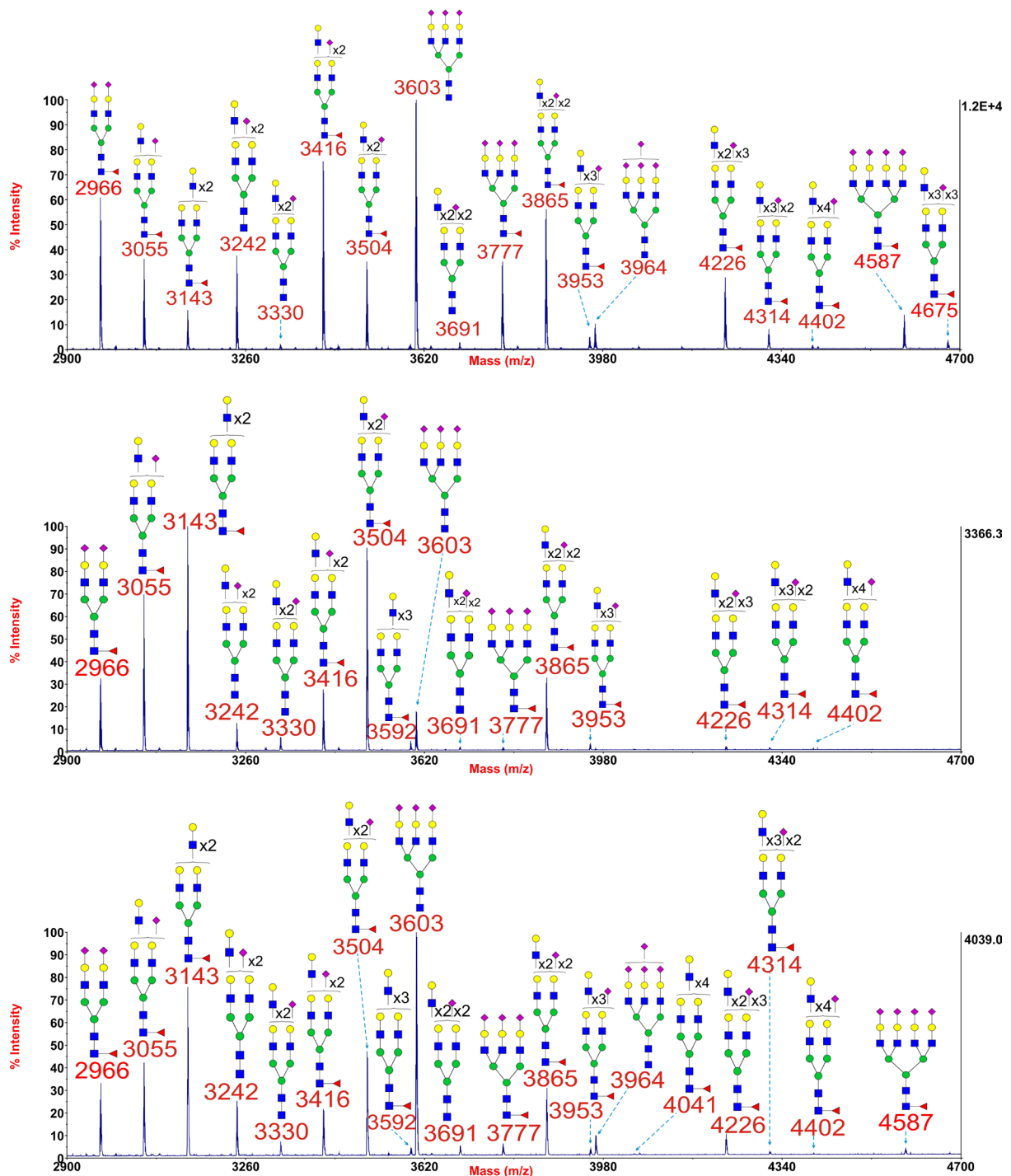
Figure 4.5 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) from myoblasts of healthy control 2 (top), *GFPT1* patient 2 (middle), the *MTND5* patient (bottom)

This figure amplifies the mass range (m/z 2900-4700) where the majority of tri- and tetra-antennary glycans are found in Figure 4.4. Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
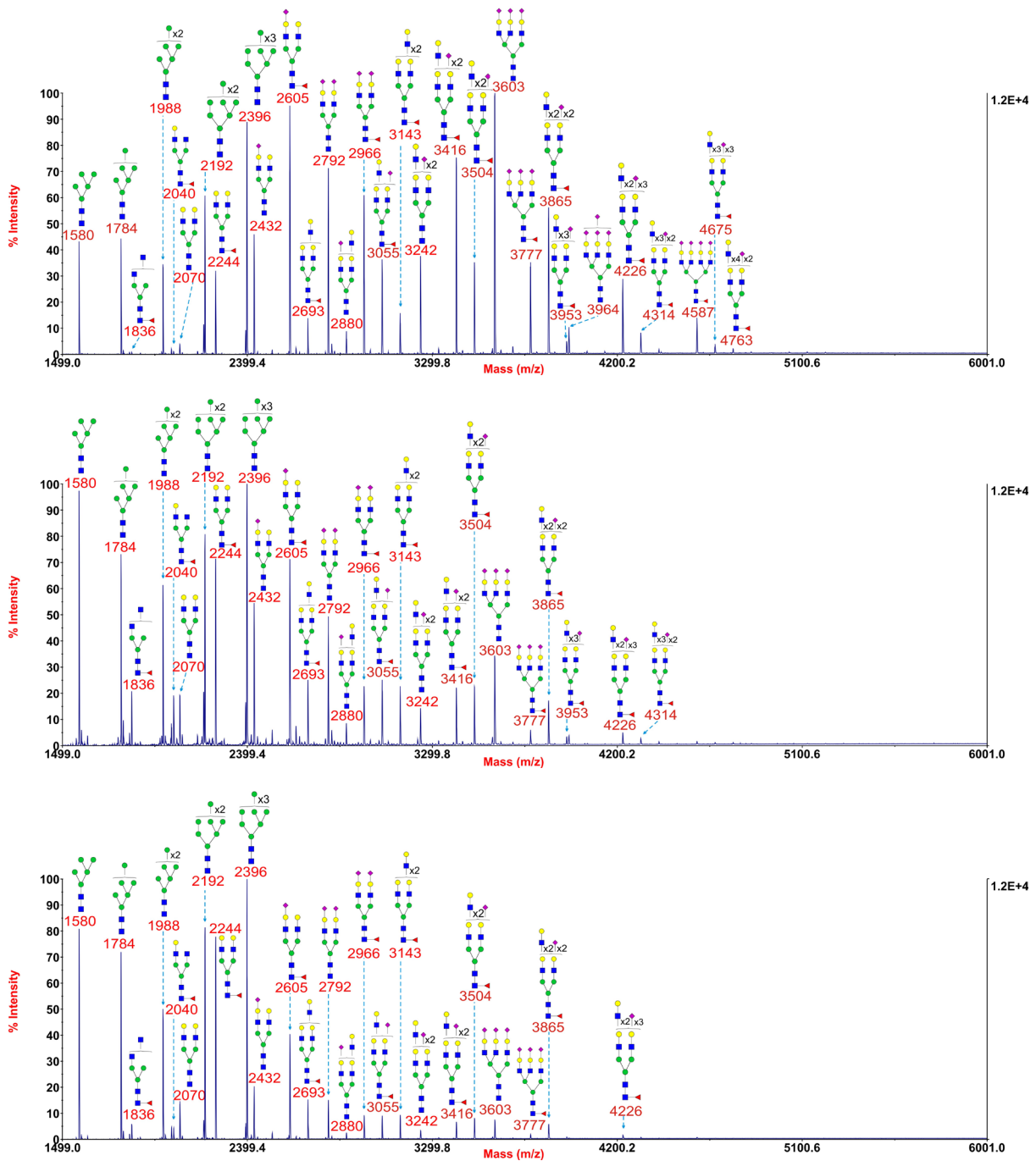
Figure 4.6 Annotated MALDI-TOF MS spectra of permethylated N-glycans from myoblasts of healthy control 2 (top), LGMD2A patient 2 (middle), Pompe disease patient (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

146
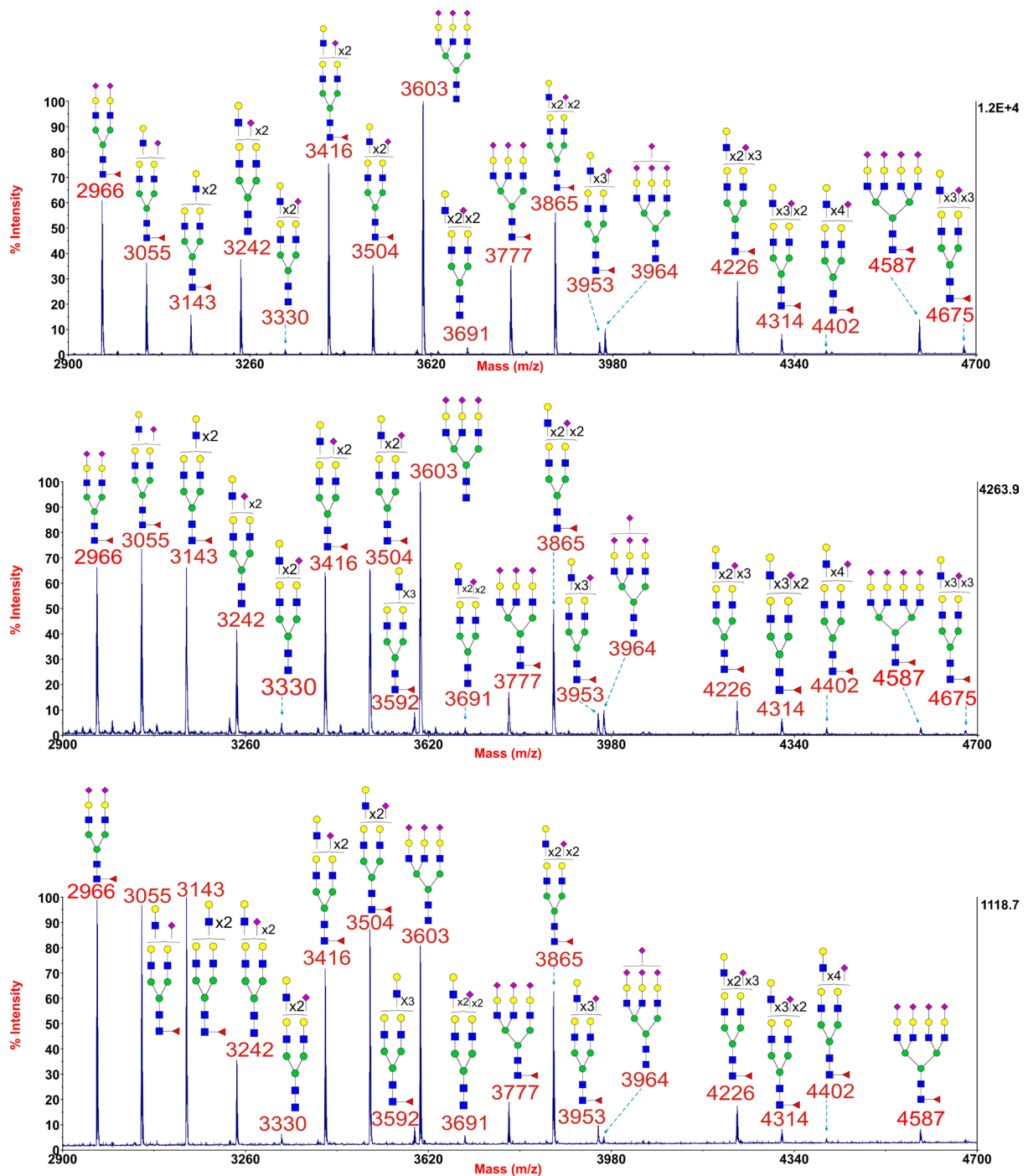
Figure 4.7 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) from myoblasts of healthy control 2 (top), LGMD2A patient 2 (middle), Pompe disease patient (bottom)

This figure amplifies the mass range (m/z 2900-4700) where the majority of tri- and tetra-antennary glycans are found in Figure 4.6. Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]⁺. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

In addition, MALDI spectra for biological replicates are shown in Figure 4.8; the result shows that there is no significant difference in the N-glycan profile of biological replicates.
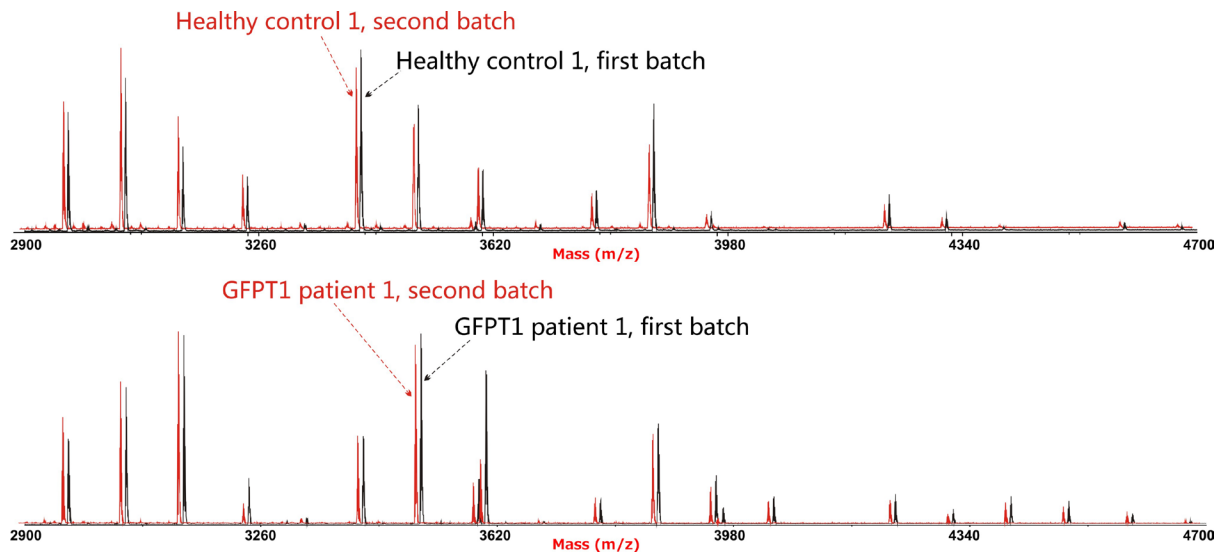
Figure 4.8 MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) from two batches of healthy control 1 myoblasts (top) and two batches of *GFPT1* patient 1 myoblasts (bottom)

See legend to Figure 3.1 for explanation of spectra arrangement. In healthy control 1, the average peak variation between the spectra from different batches is approximately 18%. In GFPT1 patient 1, the average peak variation between the spectra from different batches is approximately 14%.

To check whether the *GFPT1* patients exhibited impaired N-glycan branching, the N-glycans were classified into families with different levels of sialylation and then glycan abundances within these families were compared. Sialylation levels were determined to be similar in the various myoblast samples, which was achieved by calculating abundance ratios of pairs of glycans in which one glycan of the pair possessed one more sialic acid than the other but otherwise other compositions in the pair are identical. The results of these calculations are displayed in Figure 4.9 for the various myoblast samples. The data show that there is no significant difference in the glycan sialylation between these samples and sialylation patterns are broadly similar.
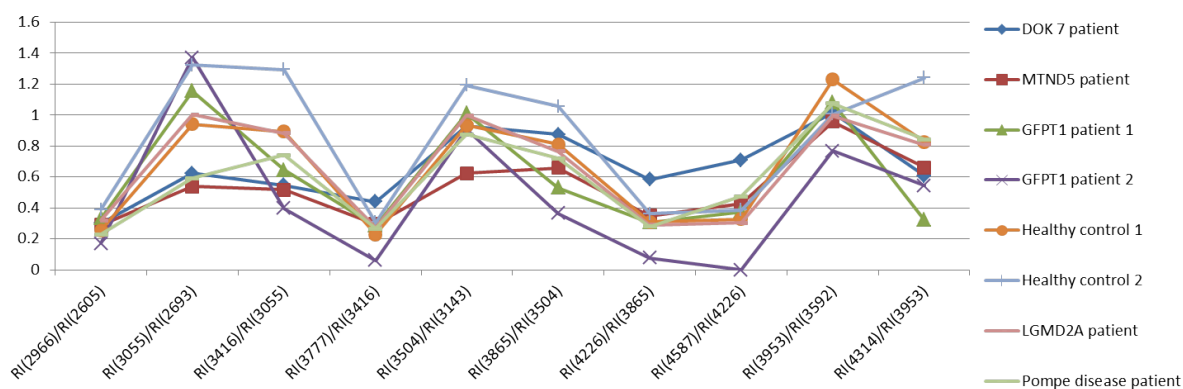
Figure 4.9 Comparison of N-glycan sialylation in the myoblasts

Each point in the graph indicates a ratio which was obtained by comparing the relative intensity (RI) of one glycan to that of another glycan which possesses one fewer NeuAc. The numbers in the brackets correspond to the m/z of the comparing glycans.

The N-glycans were classified into nonsialylated, monosialylated, disialylated, trisialylated and tetrasialylated families. For each family of glycans, the abundance ratios of pairs of glycans varying in composition by a single LacNAc unit were determined. For instance, Figure 4.10 shows comparative data for nonsialylated (Panel A) and monosialylated (Panel B) glycans possessing from two to six LacNAc units. All samples showed a similar profile of abundance ratios, albeit there is a two to three-fold divergence of ratios when comparing four LacNAcs with three LacNAcs (m/z 3143 and 2693, Panel A), with the *GFPT1* patient samples displaying somewhat higher levels of the former than observed in the other samples.

A

Comparison of the relative intensities (RI) of non-sialylated glycans

B

Comparison of the relative intensities (RI) of mono-sialylated glycans
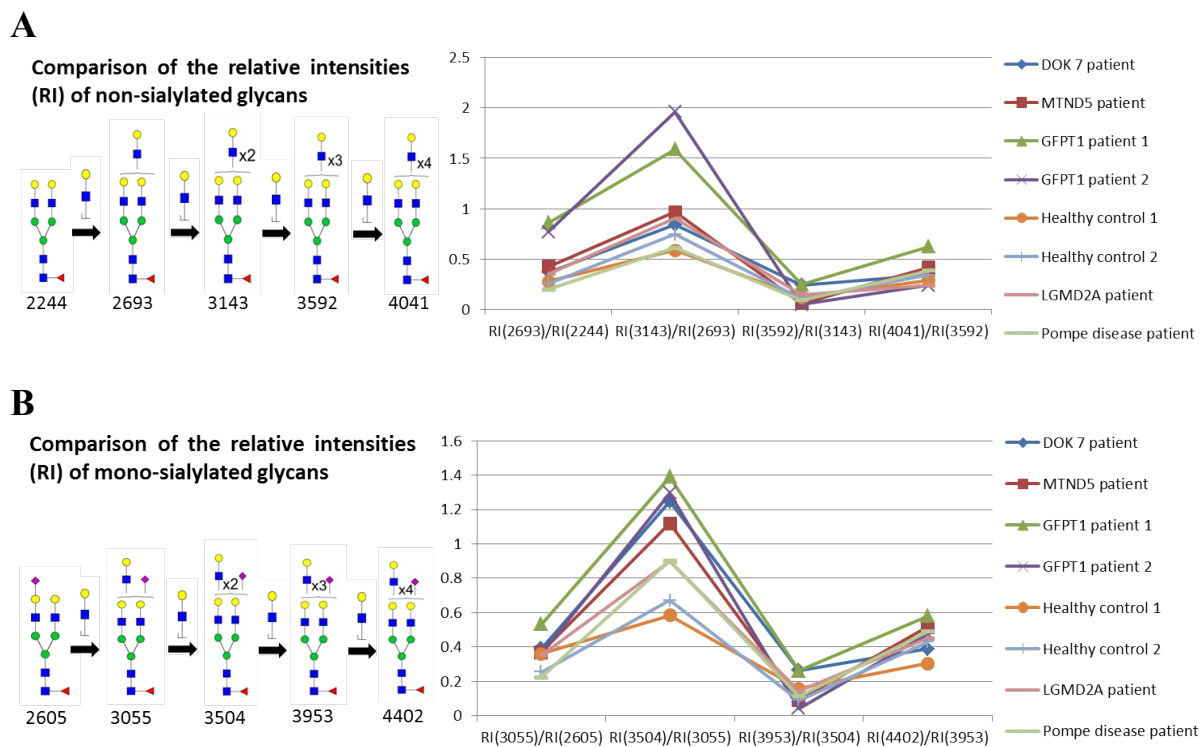
Figure 4.10 Comparison of the relative intensities of a family of nonsialylated glycans with different numbers of LacNAc (A) in myoblasts, comparison of the relative intensities of a family of monosialylated glycans with different numbers of LacNAc (B) in myoblasts

Each point in the graph indicates a ratio which was obtained by comparing the relative intensity (RI) of one glycan to that of the corresponding glycan which possesses one fewer LacNAc moiety. The number under the glycan structure is the m/z value of the glycan; the number is increasing with the addition of LacNAc moiety. In each comparison, the numbers in the brackets correspond to the m/z of the comparing glycans. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

4.3.2.2. *MALDI-TOF/TOF MS/MS analyses of the N-glycans of the myoblasts from healthy controls and GFPT1 patients and other muscular disease patients*

No significant difference between the MS spectra of the *GFPT1* patients and healthy controls and other muscular disease patients was observed, and this unexpected result could not support the initial hypothesis (see section 1.7.1) that branching might be impaired in *GFPT1* myoblasts. On the contrary, the result suggested that multiantennary glycans were likely to be slightly more abundant in *GFPT1* patients than in controls. However, it is important to realise that increasing numbers of LacNAc units does not mean increased branching because these units can be present in extended oligo-LacNAc antennae rather than as additional antennae. For instance, when a LacNAc is added to the glycan at m/z 2244, this LacNAc could be an extra antenna or an extension of an existed antenna, no matter where this LacNAc is, the m/z

value of the resulting glycan should be 2693. In this case it will be difficult to determine branching alteration via only MS analyses. Indeed the MALDI data in Figure 4.2, Figure 4.4 and Figure 4.6 confirm that the N-glycans are able to extend their antennae due to the fact that many of the glycans at high mass have more than four LacNAc moieties which are the basis of a tetra-antennary glycan. Fortunately isomeric glycans varying in oligo-LacNAc extensions and branching can be distinguished and their abundances can be compared via analysing characteristic fragment ions in MS/MS analyses. Therefore additional MS/MS investigations were carried out.

In the MS/MS analyses, it is not very accurate to compare the relative abundance of an identical glycan isoform in two different samples as the isoform may account for a greater proportion of the glycan mixture in one sample but the corresponding relative abundance may be lower than the same isoform in the other sample. Therefore glycan isoform ratios were compared within the same MS/MS analysis. Because the amounts of sample were limited, it was not practicable to characterise all components by MS/MS. Therefore only peaks that were expected to be most informative for comparative purposes were analysed with MS/MS. Note, though, that analysing every glycan peak is usually unnecessary because the N-glycans are closely related since they share a biosynthesis pathway.

For example Figure 4.11 shows MS/MS spectra acquired from the monosialylated glycan at m/z 3055 that contains three LacNAc units in the MALDI data (see Figure 4.2) for healthy control 1, *GFPT1* patient 1 and the *DOK7* patient. There are two antennae arrangements which are consistent with this composition: triantennary and biantennary with one of the antennae possessing one LacNAc extension. As shown in Figure 4.11, the MS/MS spectra are dominated by fragment ions arising from loss of a single terminal LacNAc, with or without a NeuAc. These fragment ions can be obtained from both the bi- and tri-antennary structures. Nevertheless there are several fragment ions that are characteristic for the biantennary structure with extended LacNAc. These ions are observed at m/z 935, 1781 and 2142. The peak signals of them in the spectra are weak. Importantly their relative abundances to the major fragment ions are similar in the three samples. It is estimated from abundance comparisons that the ratios of the biantennary glycan is in myoblasts from healthy control 1, *GFPT1* patient 1 and the *DOK7* patient are approximately 10.6%, 13.1% and 10.4%, respectively.
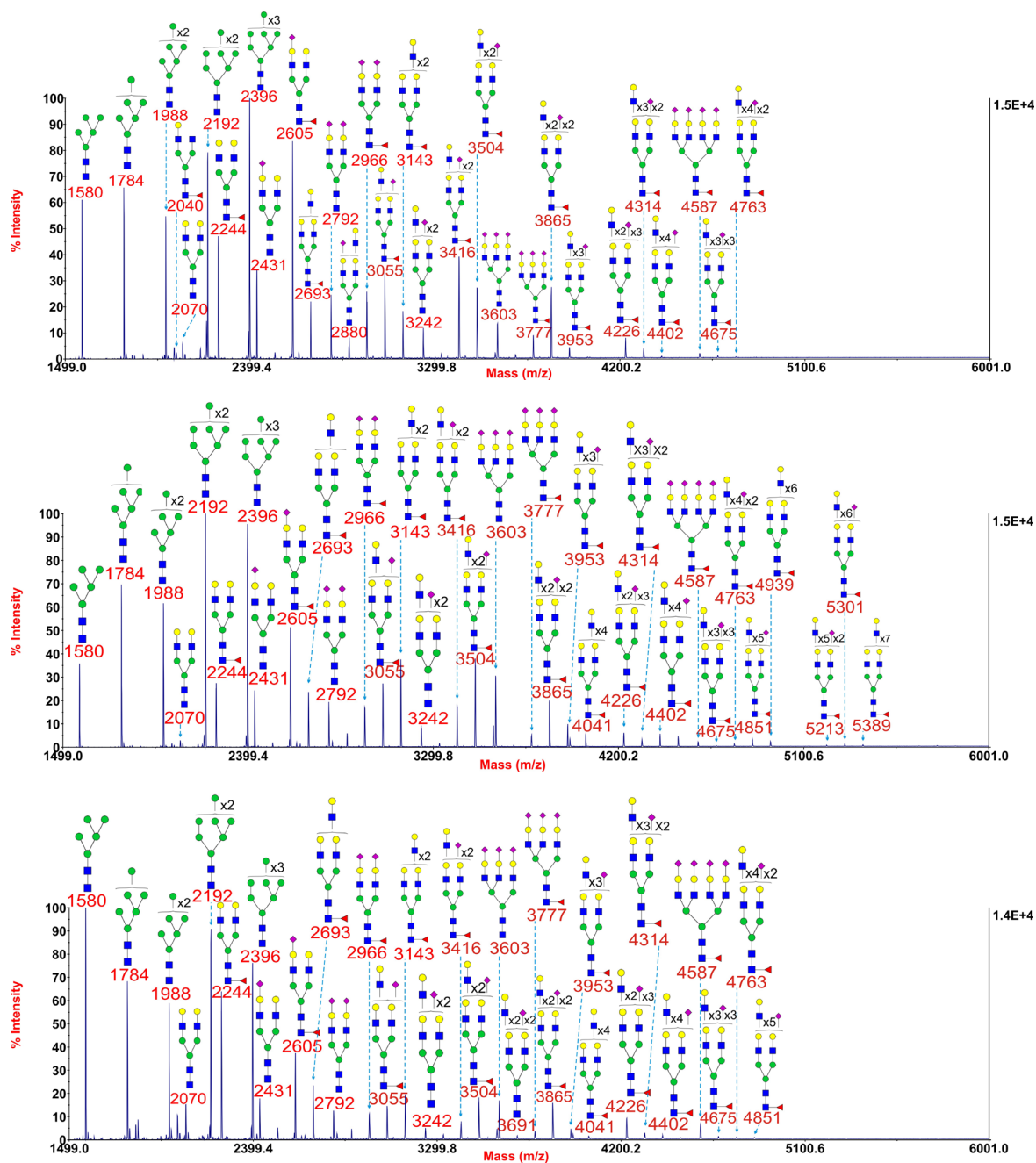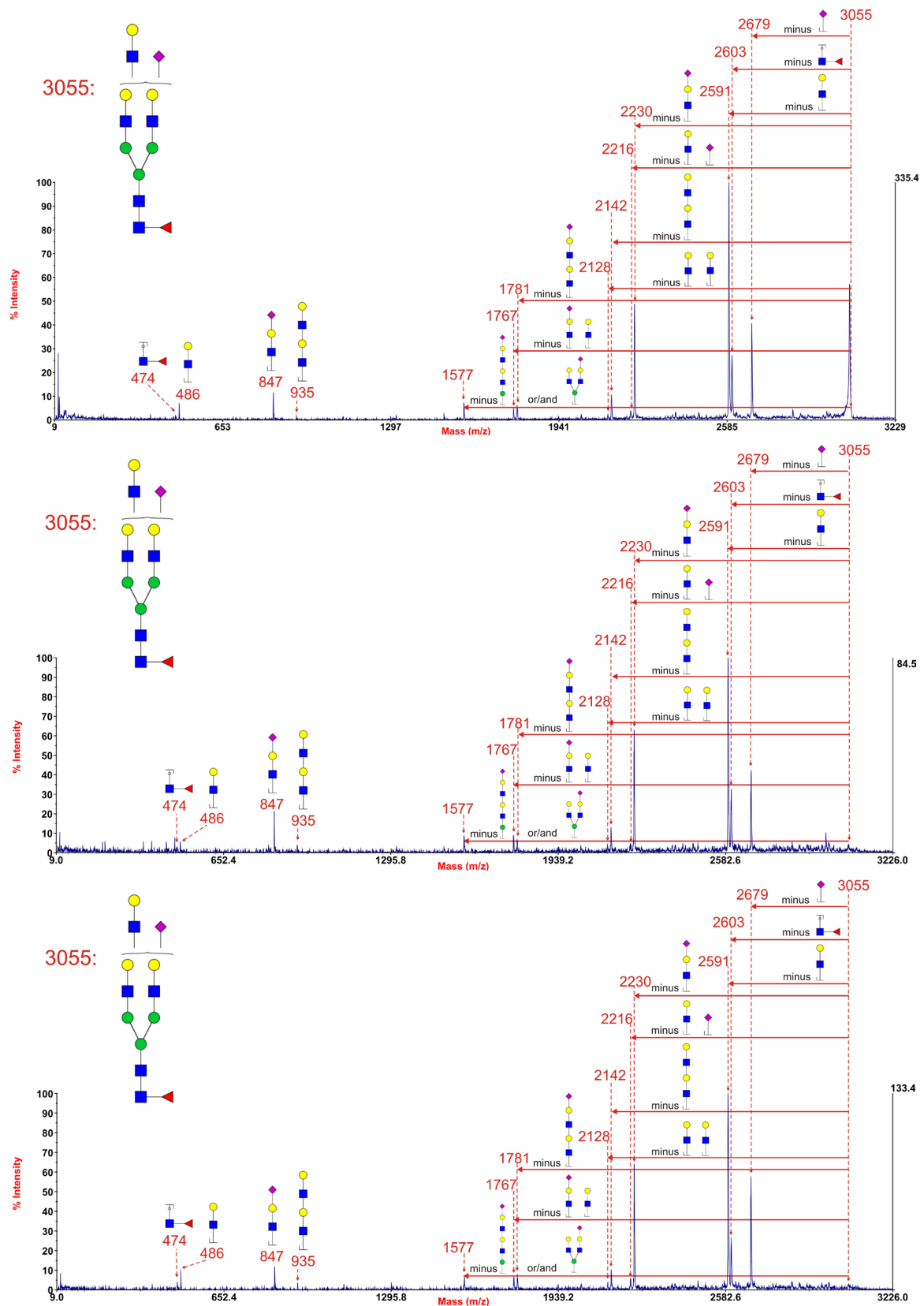
Figure 4.11 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated N-glycan at m/z 3055 in myoblasts from healthy control 1 (top), *GFPT1* patient 1 (middle) and the *DOK7* patient (bottom)

Data were acquired in the form of [M+Na]⁺ ions. See legend to Figure 3.5 for explanation of structure assignments. ■ GlcNAc, ● Man, ● Gal, ▲ Fuc, ◆ NeuAc.

The analyses of other glycans with similarly ambiguous compositions gave similar results as described above. One more example is shown in Figure 4.12 which was obtained from the glycan at m/z 2693 which contains one fewer NeuAc than the previously mentioned glycan m/z 3055 (see Figure 4.11) in the MALDI data (see Figure 4.2) for the same three subjects. There are also two antennae arragements which are consistent with this composition: triantennary and biantennary with one extended antenna. As shown in Figure 4.12, the fragment ion (m/z 2230) arising from loss of a single terminal LacNAc can be derived from both the bi- and tri- antennary structures, in addition to that, there is no characteristic ion for triantennary structure but there are two characteristic ions for the biantennary structure. These are observed at m/z 935 and 1781. The peaks of them in the spectra are also minor and their relative abundances to the major fragment ion (m/z 2230) are similar in the three samples. Similarly, the ratios of the biantennary glycan in myoblasts from healthy control 1, *GFPT1* patient 1 and the *DOK7* patient are estimated to be approximately 9.2%, 14.1% and 12.3% respectively.

Taken together, the MS and MS/MS results provide strong evidence that the patterns of glycan antennae are very similar amongst the myoblasts from healthy controls, *GFPT1* patients, and other muscular disease patients.
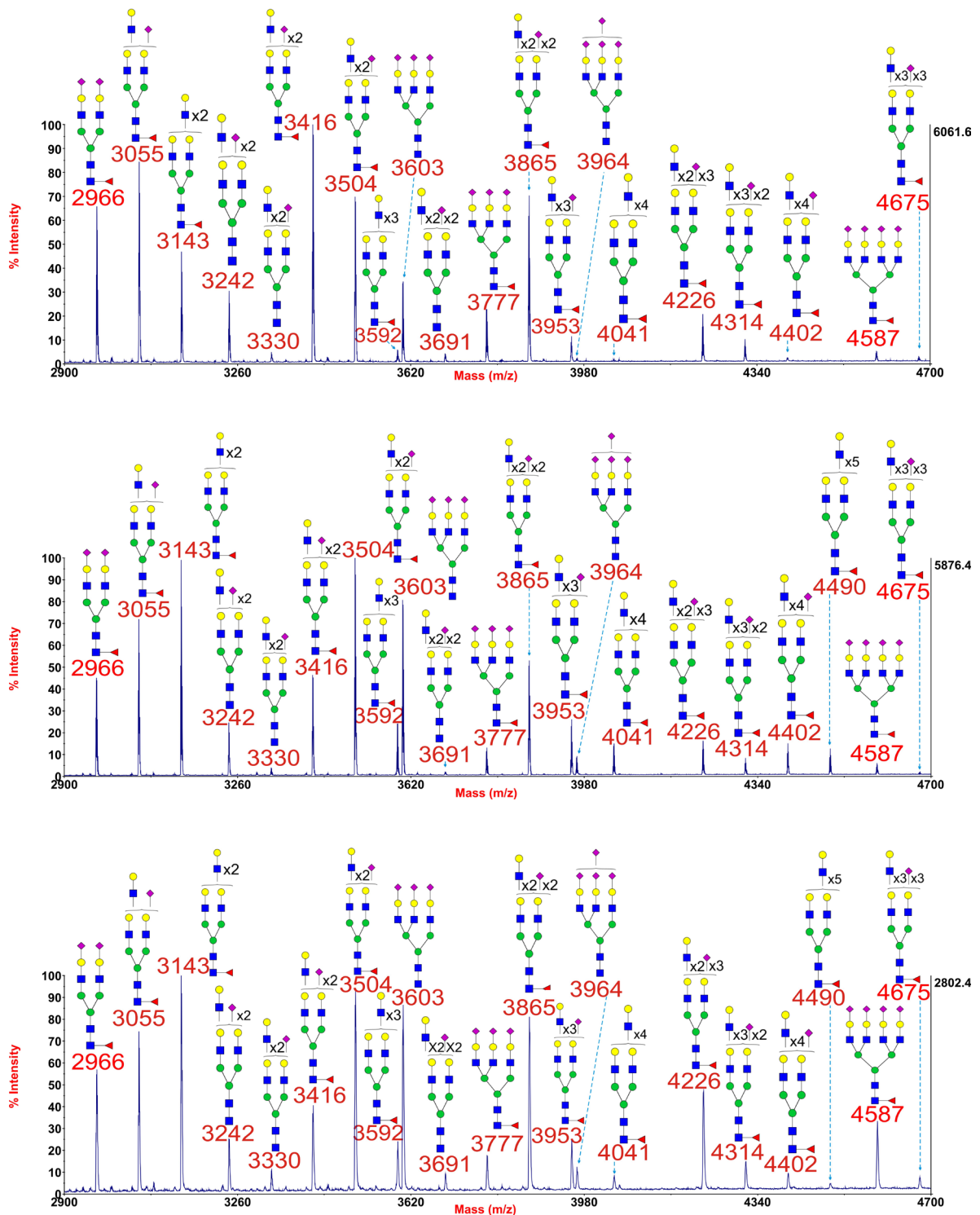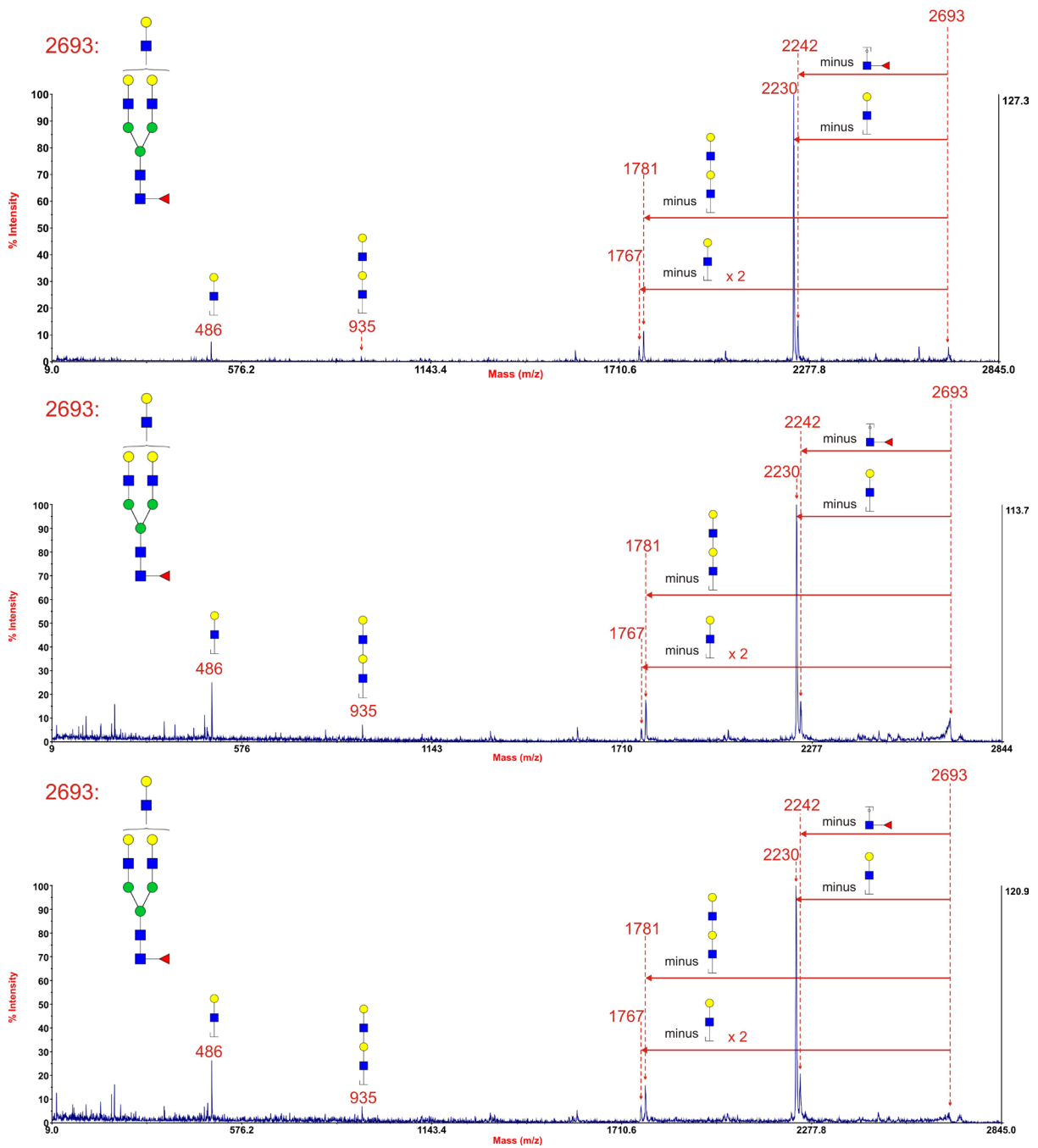
Figure 4.12 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated N-glycan at m/z 2693 in myoblasts from healthy control 1 (top), *GFPT1* patient 1 (middle) and the *DOK7* patient (bottom)

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

4.3.2.3.   Sialidase S digestion of N-glycans of the myoblasts from the DOK7 patient

The myoblast sample from *DOK7* patient in the second batch was split into two aliquots. One aliquot was used for normal glycomic analysis, and the other could be used for sialidase S digestion to determine the NeuAc linkages in myoblast N-glycome, which could give us more information concerning the myoblast N-glycome. This enzyme is specific for α2,3 linked sialic aicd.

As shown in Figure 4.13 (bottom), sialidase S digestion of N-glycans from the *DOK7* patient myoblast resulted in nearly complete desialylation of all of the core fucosylated N-glycans. A handful of sialylated glycans were still shown at m/z 2431, 2880 and 3242. These α2,6 linked sialylated glycans are lacking core fucose and are likely to be derived from FCS in the culture medium (Monk et al., 2006). Therefore, I can conclude that the NeuAc in the myoblast N-glycans is predominantly α2,3 linked.
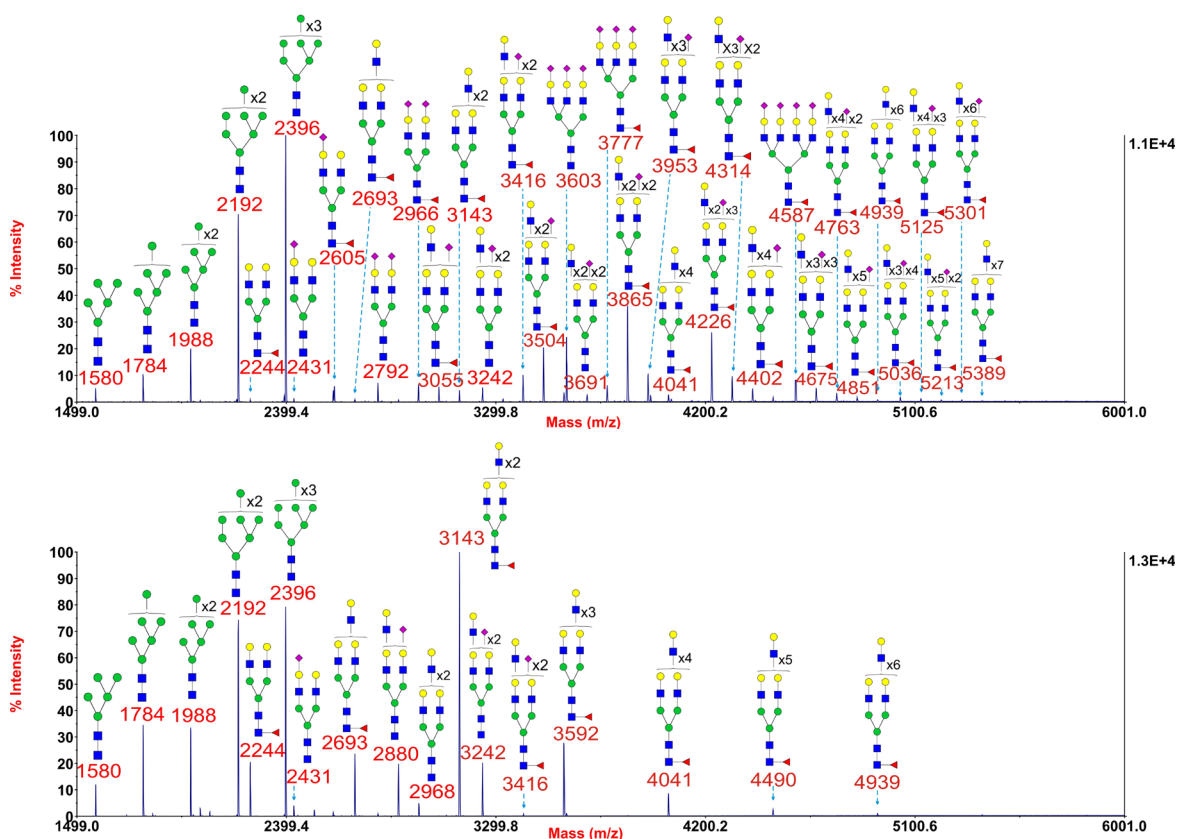
Figure 4.13 Annotated MALDI-TOF MS spectra of permethylated N-glycans (top) and sialidase S treated N-glycans (bottom) of myoblasts from *DOK7* patient, second batch

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

4.3.3. Myotube N-glycans

4.3.3.1. *MALDI-TOF MS analysis of the N-glycans of myotubes from healthy controls and GFPT1 patients and other muscular disease patients*

Myotube samples listed in Table 4.1 were analysed. These myotubes were obtained via *in vitro* differentiation of the myoblasts that used in the previous experiment. They were subjected to the same glycomic analysis that used in the myoblasts. Most of the myotube samples showed good quality data. However, the high m/z glycans were not always observed due to the fact that myotube sample quantities were more limited than for myoblasts. Representative MALDI-TOF MS spectra from healthy control 1, a *GFPT1* patient 1 and the *DOK7* patient are shown in Figure 4.14 and Figure 4.15.

Some common characteristics of mammalian cell N-glycomes were observed in Figure 4.14 and Figure 4.15 (see section 4.3.2.1 for detailed describtion). However, there is a difference in the myotube data between healthy control 1 and the *GFPT1* patient 1 and the *DOK7* patient, the difference is that although all showed minor signals for glycans with more than four LacNAcs, healthy control 1 and the *GFPT1* patient 1 showed a predominance of biantennary glycans and relatively high levels of glycans with three and four LacNAcs (Figure 4.15, top and middle panels) while the *DOK7* patient showed more abundant glycans with three and four LacNAcs (e.g. m/z 3055, 3143, 3504 and 3865) (Figure 4.15, bottom panel).
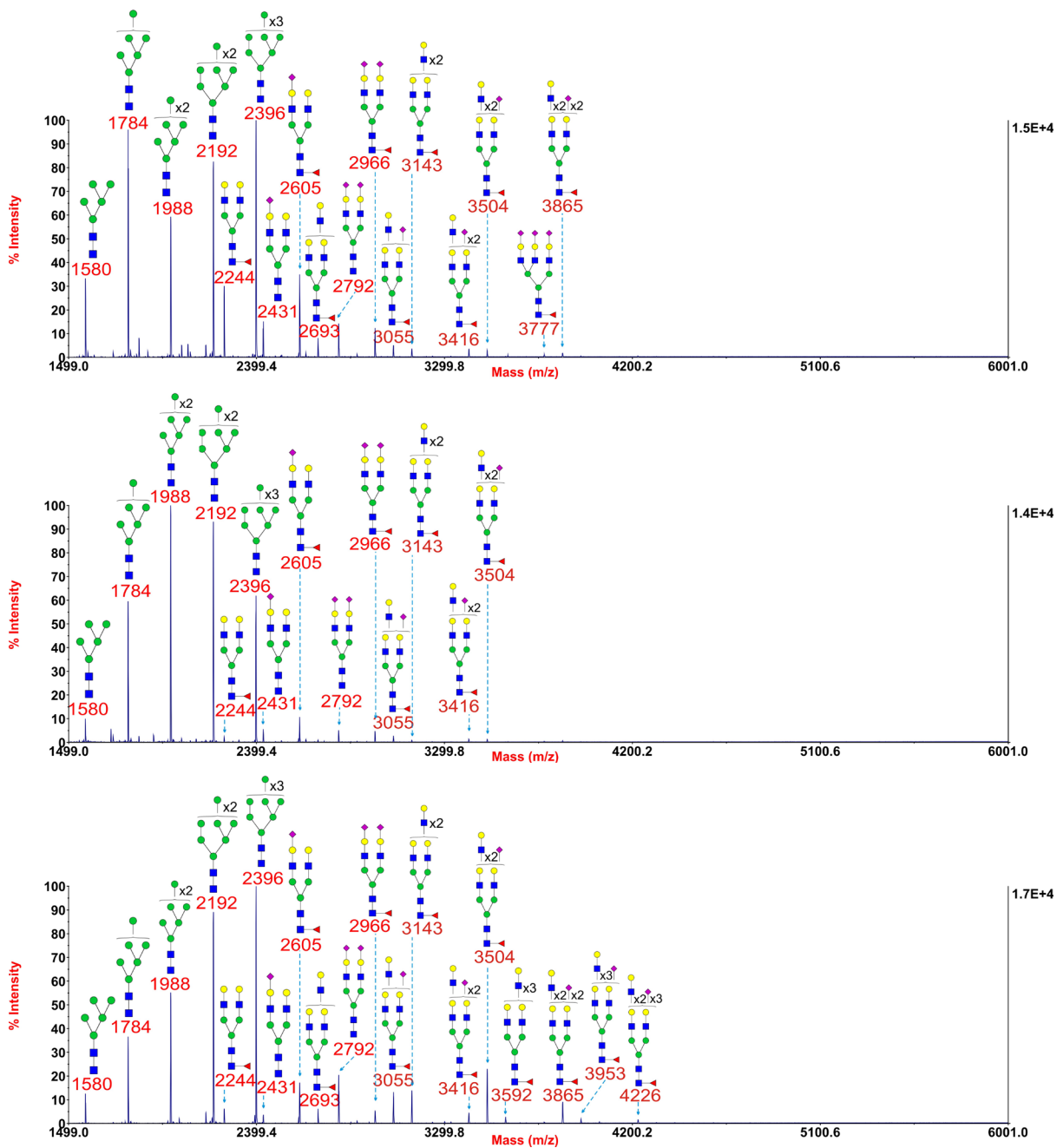
Figure 4.14 Annotated MALDI-TOF MS spectra of permethylated N-glycans of myotubes from healthy control 1 (top), *GFPT1* patient 1 (middle) and the *DOK7* patient (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
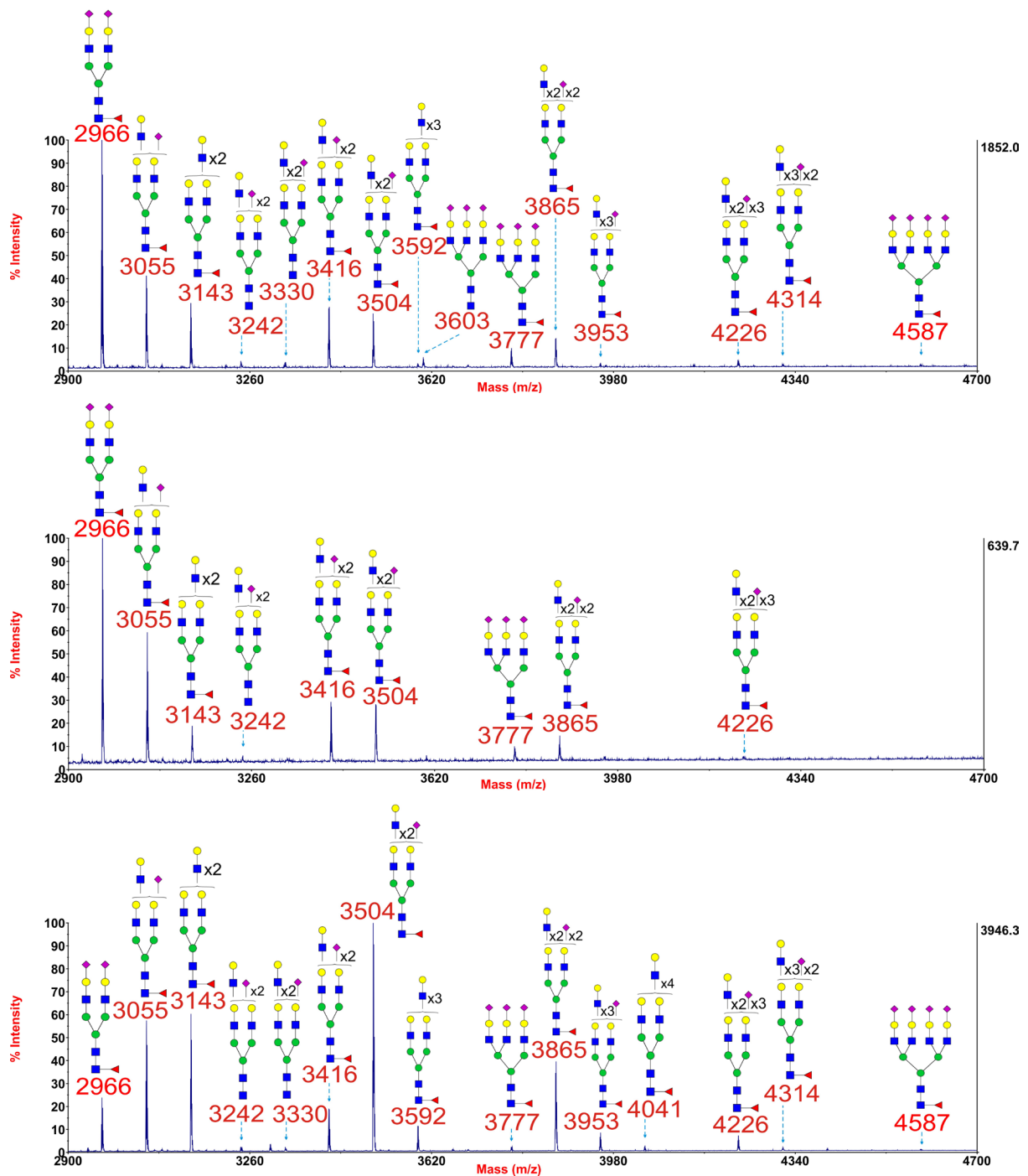
Figure 4.15 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) of myotubes from healthy control 1 (top), *GFPT1* patient 1 (middle) and the *DOK7* patient (bottom)

This figure amplifies the mass range (m/z 2900-4700) where the majority of tri- and tetra-antennary glycans are found in Figure 4.14. Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

MALDI MS data for the other patients and control are shown in Figure 4.16 to Figure 4.19; they all showed generally similar glycan profiles as those shown in healthy control 1 and the *GFPT1* patient 1 (Figure 4.14 and Figure 4.15).
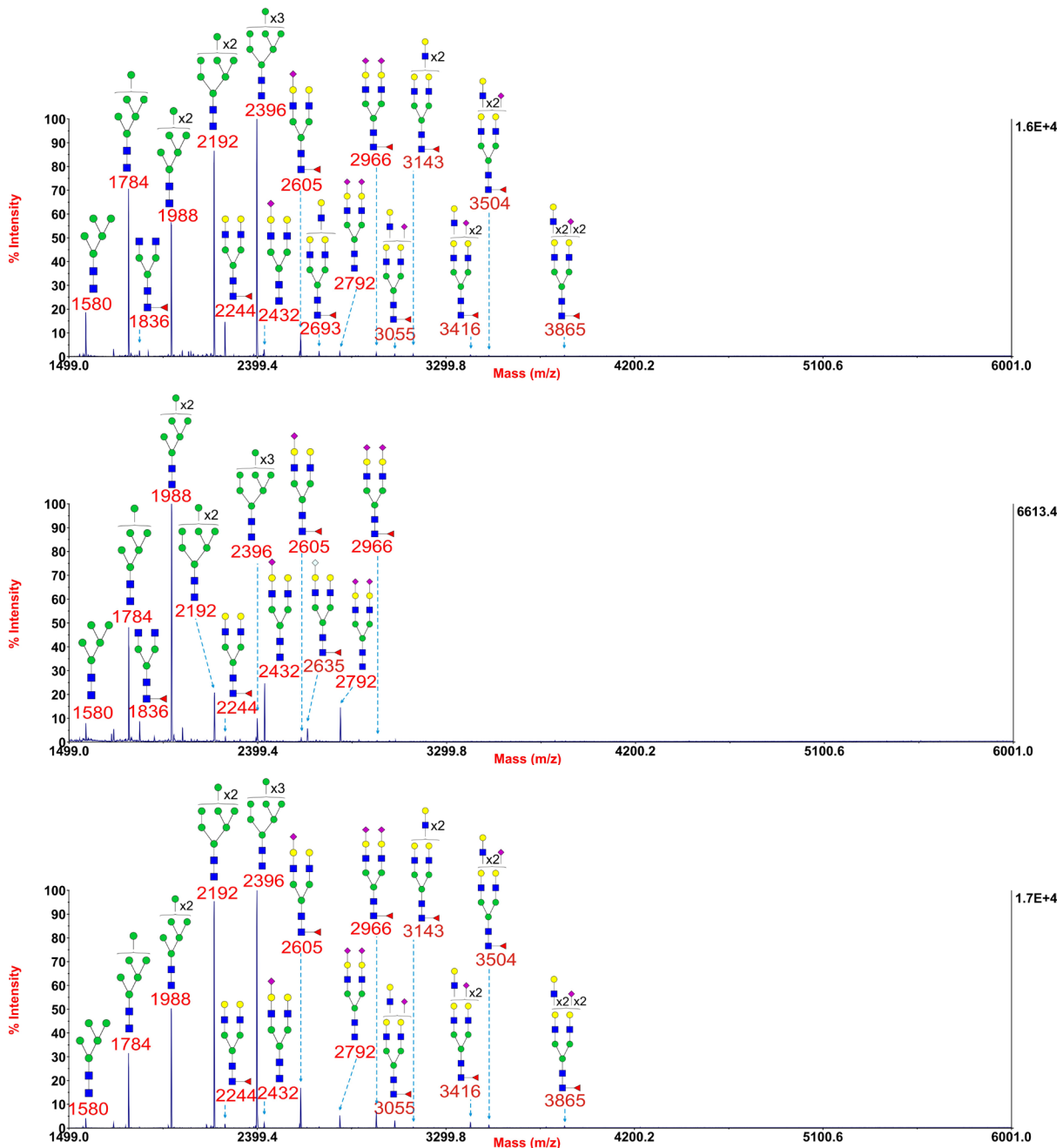
Figure 4.16 Annotated MALDI-TOF MS spectra of permethylated N-glycans from myotubes of healthy control 2 (top), *GFPT1* patient 2 (middle), the *MTND5* patient (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
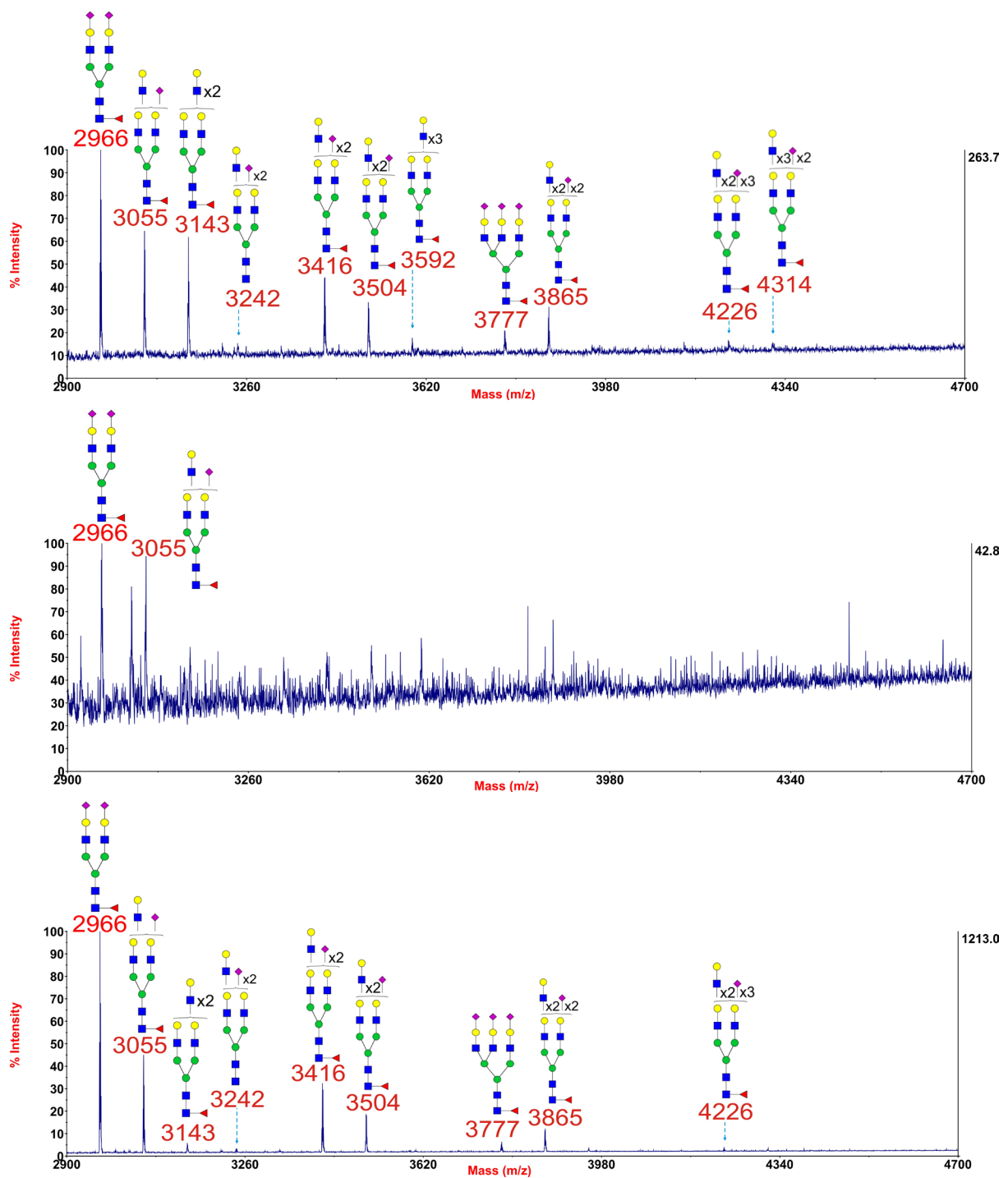
Figure 4.17 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) from myotubes of healthy control 2 (top), GFPT1 patient 2 (middle), the MTND5 patient (bottom)

This figure amplifies the mass range (m/z 2900-4700) where the majority of tri- and tetra-antennary glycans are found in Figure 4.16. Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are $[M+Na]^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
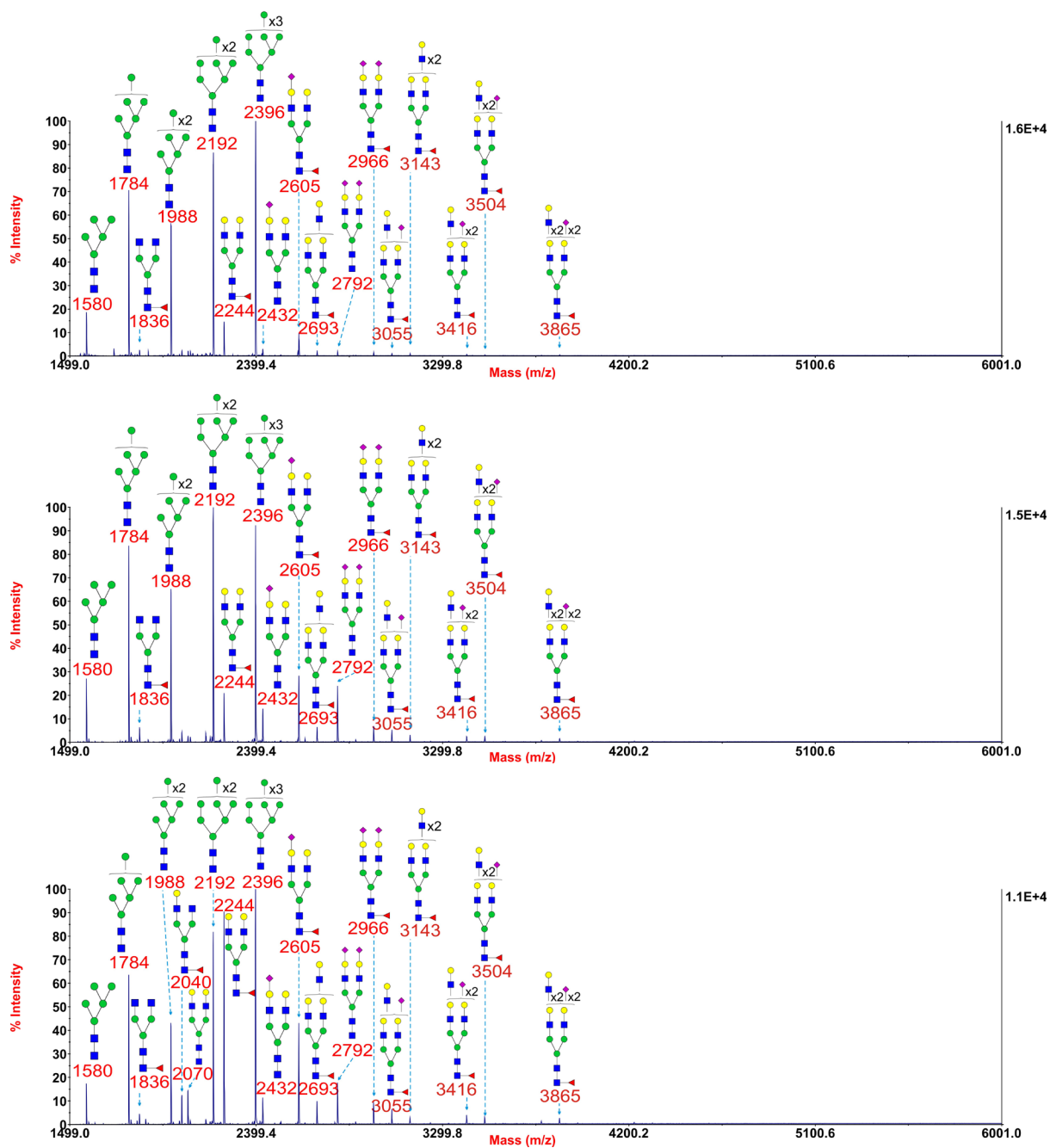
Figure 4.18 Annotated MALDI-TOF MS spectra of permethylated N-glycans from myotubes of healthy control 2 (top), LGMD2A patient 2 (middle), the MTND5 patient (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ◯ Gal, ▲ Fuc, ◆ NeuAc.
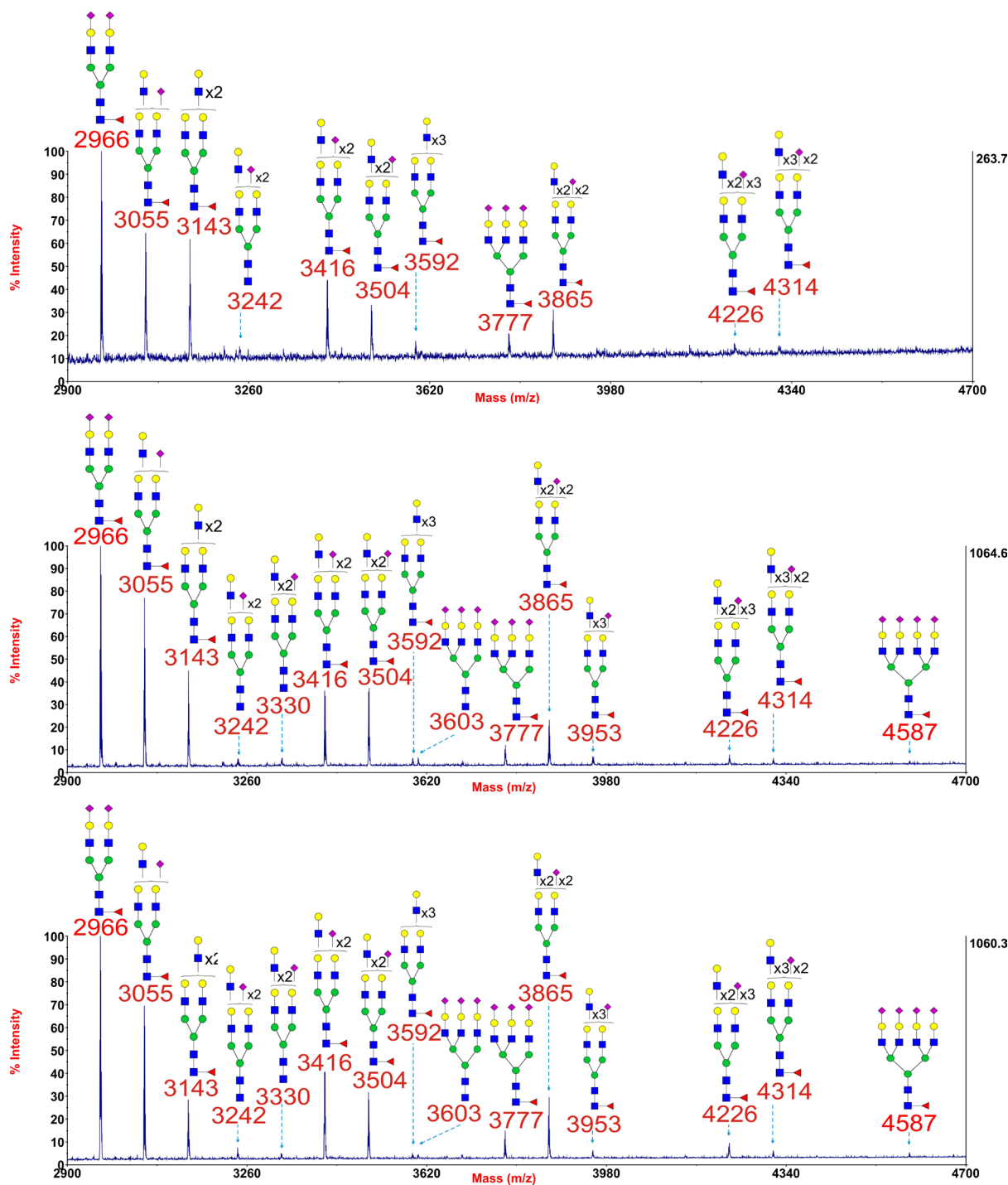
Figure 4.19 Annotated MALDI-TOF MS spectra of permethylated N-glycans (m/z 2900-4700) from myotubes of healthy control 2 (top), LGMD2A patient 2 (middle), the MTND5 patient (bottom)

This figure amplifies the mass range (m/z 2900-4700) where the majority of tri- and tetra-antennary glycans are found in Figure 4.18. Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

In order to check whether the *GFPT1* patients exhibited impaired N-glycan branching in myotubes, the N-glycans were also classified into families with different sialylation levels and then the glycan abundances within these families was compared. The sialylation levels were determined to be similar in the various myotube samples, which was achieved via using the same method employed in section 4.3.2.1. The results of these calculations are shown in Figure 4.20 for the various myotube samples. In the figure, the red line which stands for *MTND5* patient is obviously different from other lines at two comparison points (m/z 3055 and 2693, m/z 3504 and 3143), especially when comparing the abundance of the glycan with three LacNAcs and one NeuAc (m/z 3055) to the that of the glycan with three LacNAcs (m/z 2693), the resulting ratio is two to seven-fold of others. In addition, at the same comparison point (m/z 3055 and 2693) somewhat higher level of ratios (two to three-fold) are shown in *GFPT1* patient 1 and the *DOK7* patient. Except for these differences, there is no other evident difference in the glycan sialylation between the samples and sialylation patterns are broadly similar.
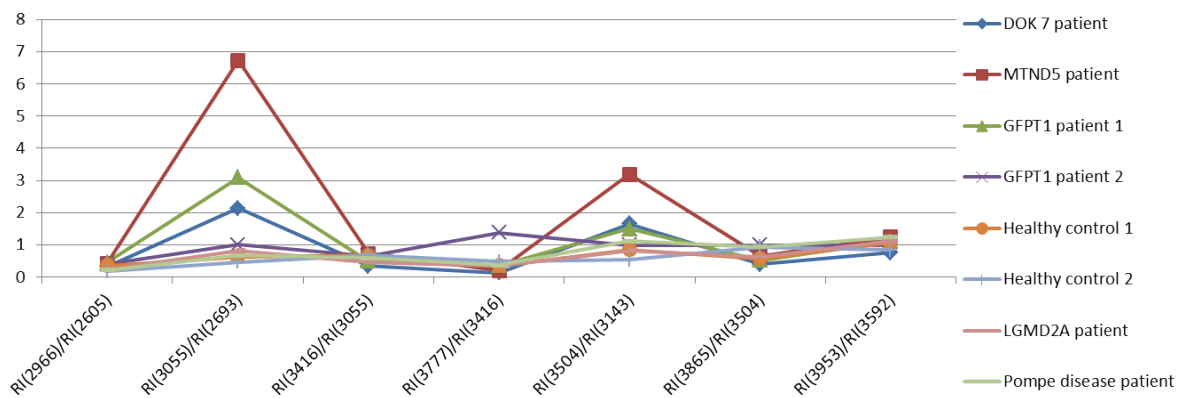
Figure 4.20 Comparison of N-glycan sialylation in the myotubes

See legend to Figure 4.9 for explanation of structure assignments.

Then the N-glycans were compared in the same way as described in section 4.3.2.1 (same sialylation level, but different number of LacNAc). Figure 4.21 shows myotube comparative data for nonsialylated (Panel A) and monosialylated (Panel B) glycans containing from two to six LacNAc units. The result is similar to that of myobolasts which has been described in section 4.3.2.1, except for a two to four-fold divergence of ratios when comparing 4 LacNAcs with 3 LacNAcs (m/z 3143 and 2693 in Panel A, m/z 3504 and 3055 in Panel B) shown in the *DOK7* patient.
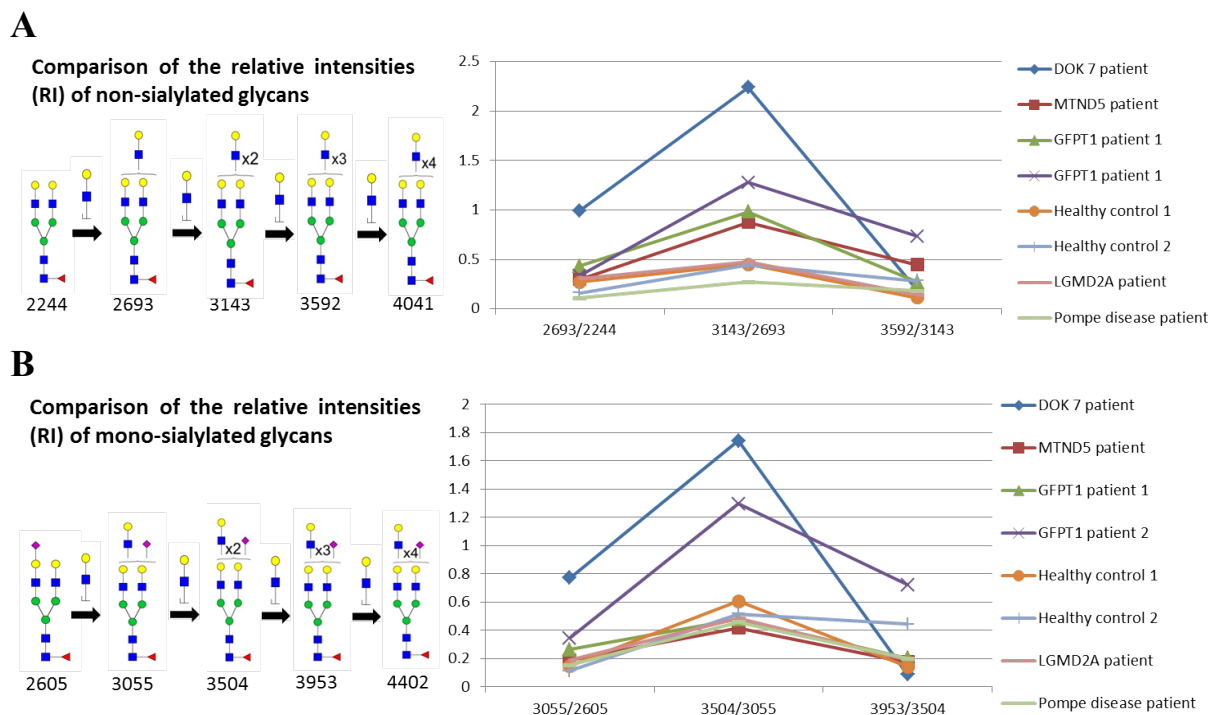
Figure 4.21 Comparison of the relative intensities of a family of nonsialylated glycans with different numbers of LacNAc (A) in myotubes, comparison of the relative intensities of a family of monosialylated glycans with different numbers of LacNAc (B) in myotubes

See legend to Figure 4.10 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

Except for the differences mentioned previously, there was no significant difference between the myotube MS spectra of the *GFPT1* patients and healthy controls and other muscular disease patients, and this unexpected result was similar to the myoblast result described in section 4.3.2.1.

4.3.3.2.   *Sialidase S digestion of N-glycans of the myotubes from the DOK7 patient*

The myotube sample from *DOK7* patient was split into two aliquots. One aliquot was used in section 4.3.3.1, and the other was used for sialidase S digestion to determine the NeuAc linkages in myotube N-glycome.

Digestion of N-glycans from *DOK7* patient myotubes with sialidase S resulted in a nearly complete desialylation on all core fucosylated N-glycans (Figure 4.22). Several minor sialylated glycans were observed at m/z 2431 and 2792. These sialylated glycans are

probably α2,6 linked sialylated glycans and they are likely to be derived from FCS (see section 4.3.2.3). Therefore, I conclude that the NeuAc linkage in myotubes is not changed upon differentiation.
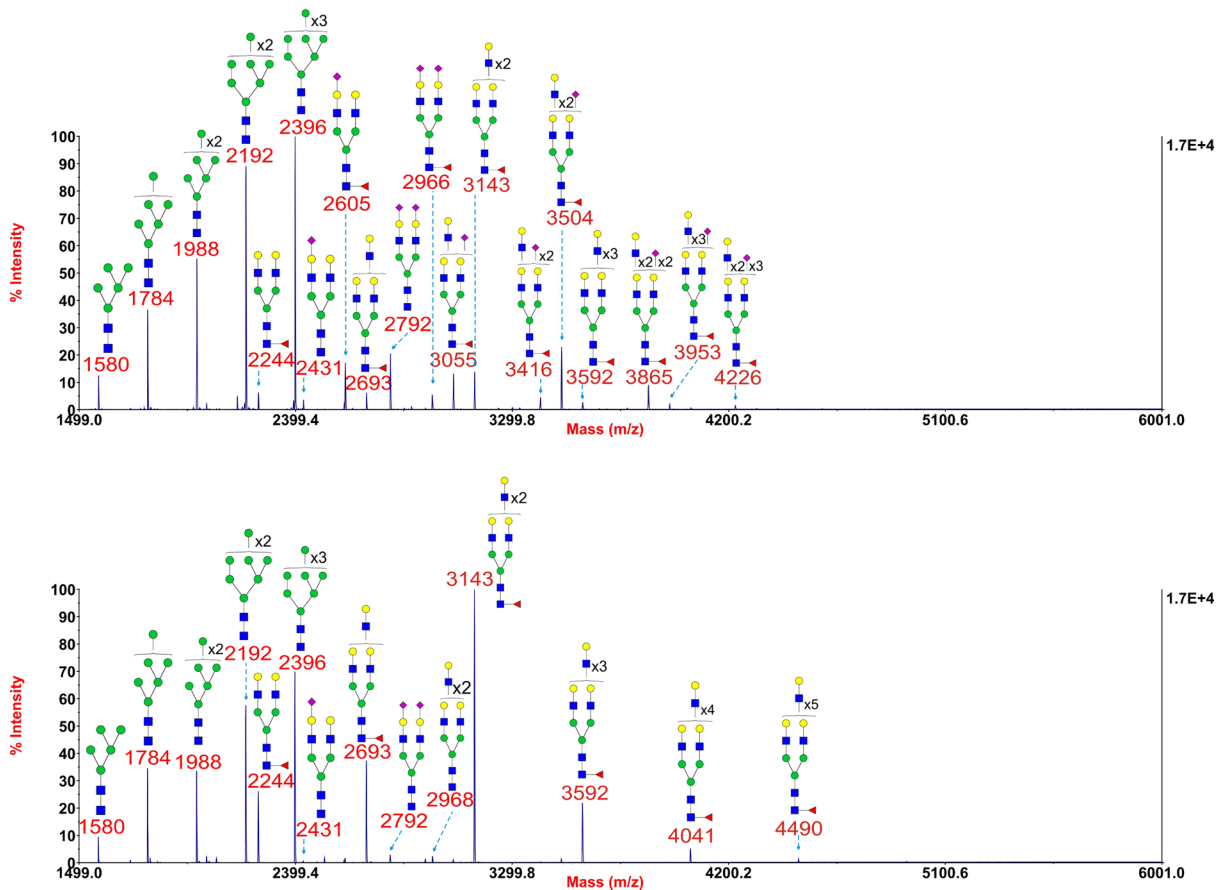
Figure 4.22 Annotated MALDI-TOF MS spectra of permethylated N-glycans (top) and sialidase S treated N-glycans (bottom) of myotubes from *DOK7* patient

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 4.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

4.3.4. Myotube O-glycans

4.3.4.1. *MALDI-TOF MS analysis of the O-glycans of the myotubes*

In addition to analysing N-glycans, the analysis of O-glycans was also performed; however, MALDI MS data were only obtained from myotubes of 2 healthy controls and 4 muscular disease patients. The O-glycan profiles are shown in Figure 4.23 and Figure 4.24. The figure showed the presence of both core 1 (m/z 534, 895 and 1257) and core 2 (m/z 779, 984, 1345

and 1706) structures. The profiles of healthy controls, the *DOK7* patient, *MTND5* patient, LGMD2A patient and Pompe disease patient are broadly similar. This result was obtained via comparing the relative abundance of glycans within the same family in each sample. In the core 1 containing glycan family, all samples showed that the relative abundance of nonsialylated glycan (m/z 534) was higher than monosialylated (m/z 895) which was higher than disialylated (m/z 1257). In the core 2 containing glycan family, all samples showed that there was no significant difference between the relative abundance of the nonsialylated glycan (m/z 984) and the monosialylated (m/z 1345).
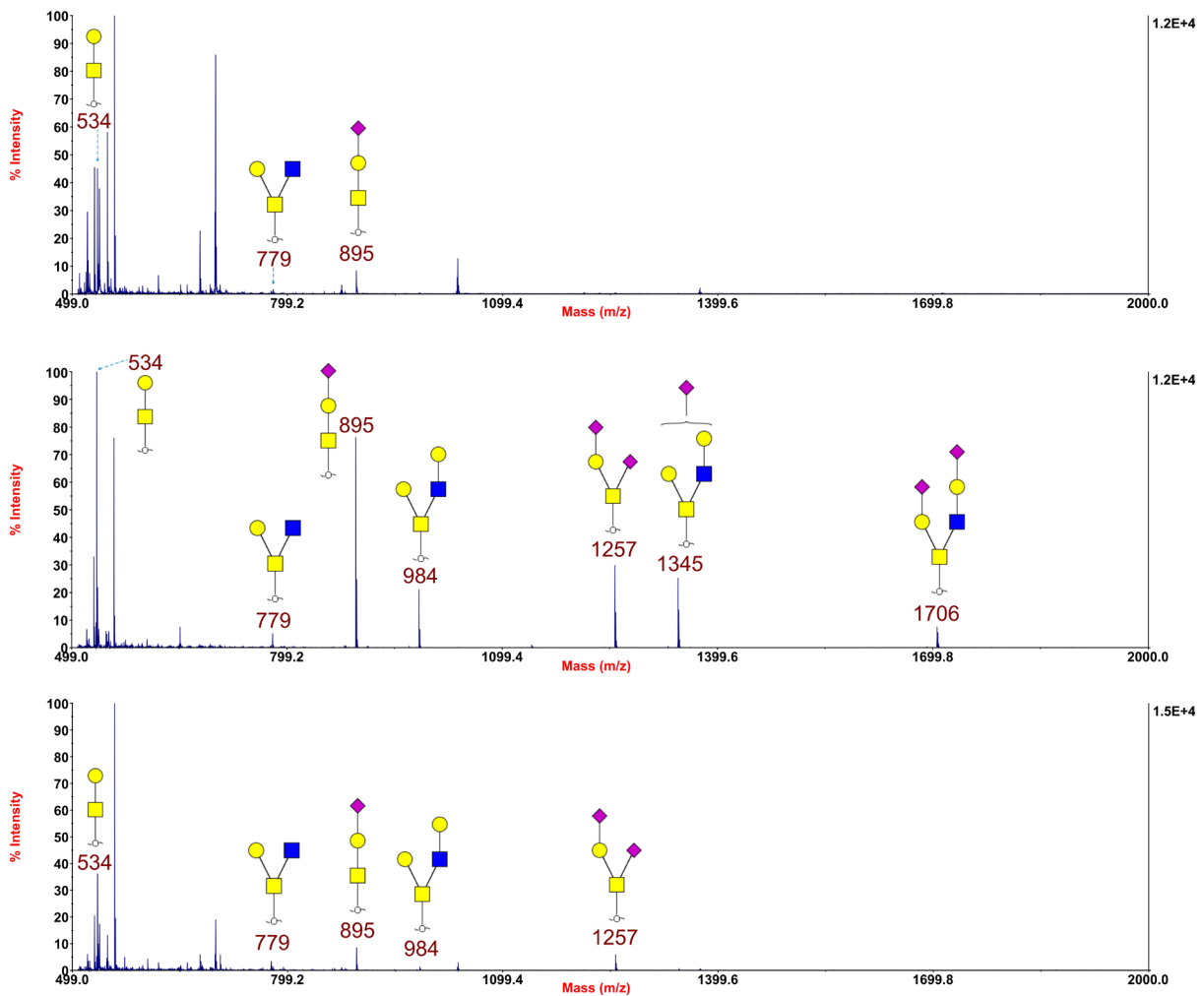
Figure 4.23 Annotated MALDI-TOF MS spectra of permethylated O-glycans of myotubes from healthy control 1 (top), the *DOK7* patient (middle), *MTND5* patient (bottom)

These profiles were obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column. All ions were obtained in the form of [M+Na]$^+$. See legend to Figure 3.13 for explanation of structure assignments. ☐ GalNAc, ■ GlcNAc, ○ Gal, ◆ NeuAc.
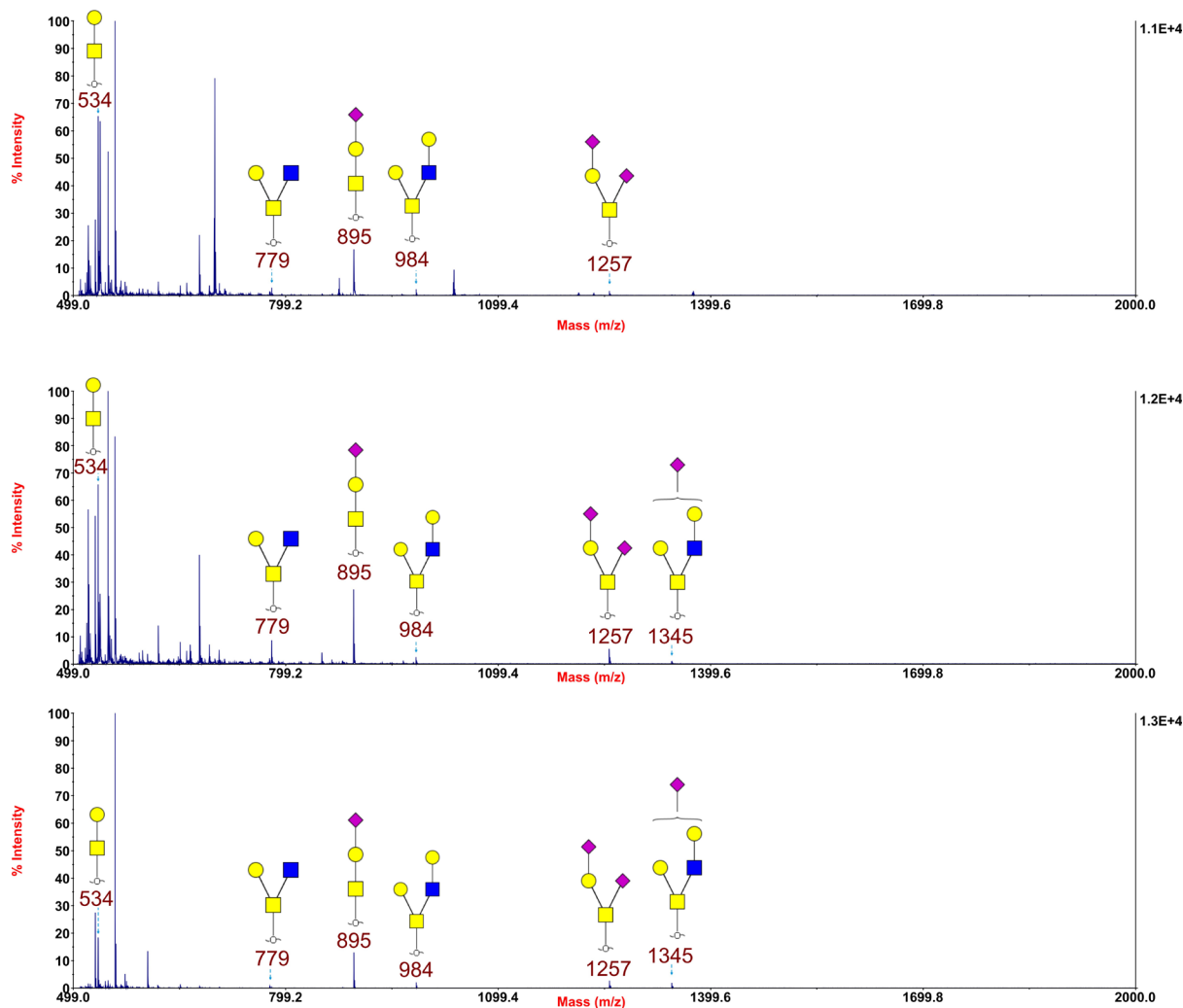
Figure 4.24 Annotated MALDI-TOF MS spectra of permethylated O-glycans of myotubes from healthy control 2 (top), LGMD2A patient (middle) and Pompe disease patient (bottom)

These profiles were obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column. All ions were obtained in the form of [M+Na]$^+$. See legend to Figure 3.13 for explanation of structure assignments. □ GalNAc, ■ GlcNAc, ○ Gal, ◆ NeuAc.

According to the O-glycan biosynthesis pathway, all O-glycan structures displayed in Figure 4.23 and Figure 4.24 are unambiguous except the structure of the glycan at m/z 895 and 1345. The MS/MS spectra of these glycans gave same results as described in section 3.3.2.1. Glycan cartoons were assigned accordingly.

Sialidase S digestion was also carried out on O-glycans to determine the NeuAc linkage. As shown in Figure 4.25 (bottom), digestion of O-glycans from *DOK7* patient myotubes with sialidase S resulted in a nearly complete desialylation on O-glycans, and with the sialylated O-glycan at m/z 895 being observed. This indicated that the NeuAc in myotubes is mainly α2,3 linked.
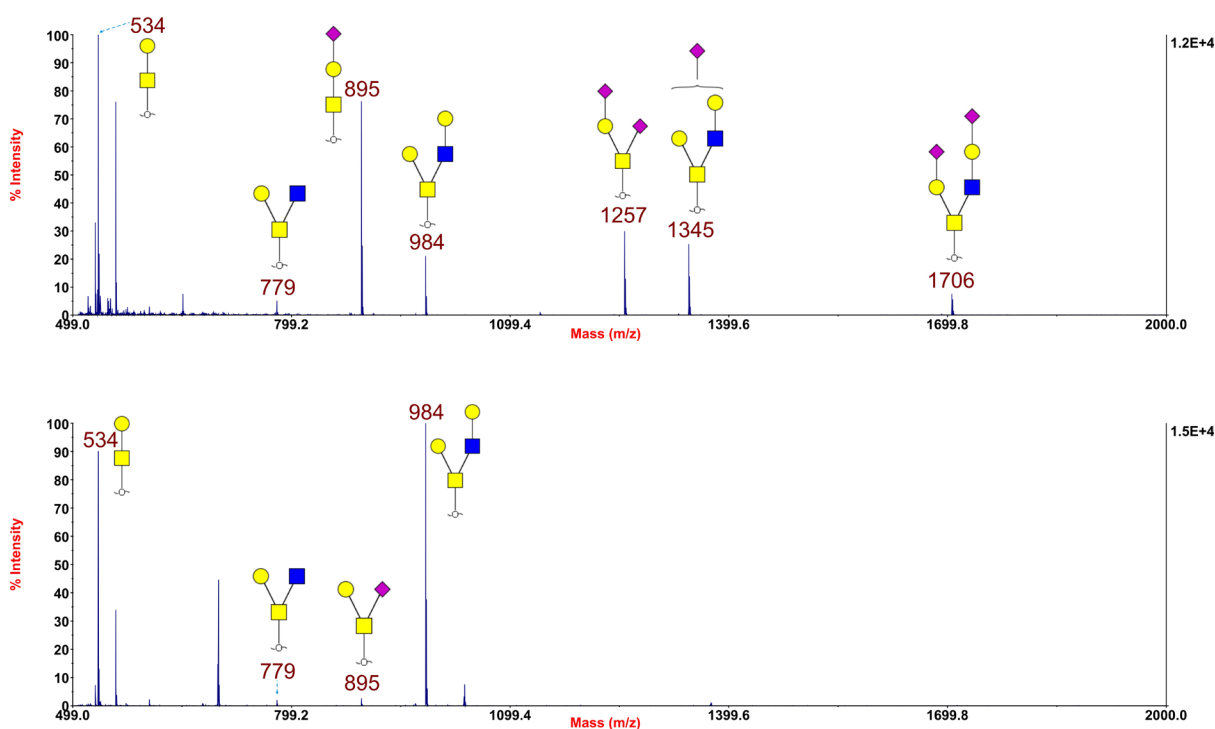
Figure 4.25 Annotated MALDI-TOF MS spectra of permethylated O-glycans (top) and sialidase S treated O-glycans (bottom) of myotubes from *DOK7* patient

Profiles were obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.13 for explanation of structure assignments. ▢ GalNAc, ◼ GlcNAc, ⬤ Gal, ◆ NeuAc.

4.4.   Discussion

It has been reported that CMS patients with *GFPT1* mutations shared clinical features with CMS patients with *DPAGT1* mutations (Belaya et al., 2012; Guergueltcheva et al., 2012). In addition, both *GFPT1* and *DPAGT1* play important roles in N-glycan biosynthesis pathway, and their mutations could result in muscle endplate AChR instability (Zoltowska et al., 2013;

Belaya et al., 2012; Freeze et al., 2015). These indicate faulty N-glycosylation of the receptors. Muscle specimens from CMS patients with *GFPT1* mutations showed that their endplates were poorly developed, ans this implied that the endplate specific glycoproteins, such as MuSK and dystroglycans, could be hypoglycosylated (Engel et al., 2015; Selcen et al., 2013). This thesis for the first time reports N-glycan profiles of *in vitro* cultured human myoblasts and N- and O-glycan profiles of human myotubes that differentiate from myoblasts *in vitro*.

Employing ultrasensitive MALDI-TOF MS combined with MALDI TOF/TOF MS/MS, I observed that more high mass glycans were detected when the cells were grown in the medium containing FCS with higher concentration (see section 4.3.1.1). FCS possesses N-glycans (Monk et al., 2006), however, their N-glycans which are non-core fucosylated can be distinguished from the ones I am interested in which are core fucosylated (see section 4.3.2). More importantly, in this project the FCS derived glycans do not seem to be a problem as they are not dominating the spectra, thus the contamination will not have too much influence on the conclusion that I made. In all the comparisons carried out in this project, the non-core fucosylated glycans are excluded.

The data of cells from *GFPT1* patient 3 gave less information (only high mannose glycans and three complex N-glycans were detected) than the other patients. This is in accordance with the fact that the cells grew more slowly than the others (see section 2.1.3.2).

The main findings of this project are: the N-glycan profiles of myoblasts (Figure 4.2 and Figure 4.4) and myotubes (Figure 4.14 and Figure 4.16) from healthy controls, *GFPT1* patients and other patients are broadly similar, the O-glycan profiles of myotubes seem to be broadly similar as well, the absence of core 2 containing glycans in healthy control 1 is probably caused by the smaller amount of the analyte.

Sialidase S digestion demonstrated the presence of α2,3 linked NeuAc in both myoblast and myotubes from *DOK7* patient; and this implies that the NeuAc in the myoblasts and myotubes from healthy controls and other patients is probably also α2,3 linked (Figure 4.13, Figure 4.22 and Figure 4.25). Indeed, selective expression of sialyltransferase is common, for instance, Galβ1,3(4)GlcNAcα2,3-sialyltransferase is specifically expressed in human K562 leukemia cells (Kitagawa and Paulson, 1994). Therefore I can reasonably conclude that my data indicate 2-3 sialyl transferases, not 2-6, are mainly expressed in myoblasts and myotubes. This could be tested using SNA-I which is an important tool in glycoconjugate

research as it possesses binding specificity to α2,6 linked NeuAc not α2,3 linked NeuAc (Shang and Van Damme, 2014; Shibuya et al., 1987). SNA-I will probably not bind or show low binding affinity to myoblasts and myotubes.

The differences observed in the *DOK7* patient and *MTND5* patient between other samples in Figure 4.20 and Figure 4.21 respectively implied that these two patients did not appear to be the same as the others and so perhaps they are not good "controls" for *GFPT1* patients and healthy controls. Another explanation for the difference is that it might be caused by individual differences between the samples, even biological replicates could show difference (see Figure 4.8). Except for these differences, there was no other difference in the comparison of the relative abundances of the bi-, tri- and tetraantennary N-glycans in all myoblast and myotube samples. These findings are unexpected as the original hypothesis (see section 1.7.1) that *GFPT1* mutations in the cells would result in changes in the N-glycan branching. This hypothesis was made based on a recent glycomic analysis concerning N-glycosylation of leukocytes from patients with mutations in the *PGM3* gene which encodes a protein immediately after GFPT1 in the hexosamine biosynthesis pathway (Sassi et al., 2014). A reduced level of tri- and tetraantennary N-glycans was observed, and this was rationalized by the fact that the initiations of four antennae are catalysed by 4 different enzymes which are GlcNAc-T I (medial Golgi-branching N-acetylglucosamine transferase I), GlcNAc-T II, GlcNAc-T IV and GlcNAc-T V, the K_m values of which are 0.04, 0.96, 5.0 and 11.0 mM UDP-GlcNAc respectively (Lau et al., 2007). In Michaelis-Menten kinetics, the value of K_m is a measure of the substrate concentration needed for effective reaction to occur and is an important characteristic of enzyme-substrate interactions (Berg et al., 2002a). Higher K_m value corresponds to lower substrate affinity and thus a lower catalytic activity. Therefore, by requiring higher substrate concentration the tri- and tetraantennary N-glycans will be most impaired by a reduction in UDP-GlcNAc concentration. However, the glycomic results did not support this hypothesis.

Although the glycomic result of the myoblasts and myotubes implies that the gene (*GFPT1*) mutations do not result in global glycan alterations, it does not rule out the possibility that individual glycoproteins have altered glycosylation. It could be that alterations in N-glycosylation are subtle and such subtle differences are lost in the general glycomic profiling or are difficult to identify. Additionally, it could be that the original glycosylation in the patient muscle is different from that in the *in vitro* cultured cells or the glycosylation might have changed during the muscle tissue regeneration. Therefore, detailed glycoproteomic

analysis of specific glycoproteins, such as acetylcholine receptor and agrin, could be carried out as future work. Indeed phenotypic similarities between CMS caused by *GFPT1* mutations and CMS caused by *DPAGT1* mutations would suggest that reduced endplate acetylcholine receptor in *GFPT1* CMS patients might at least partially be due to abnormal glycosylation (Zoltowska et al., 2013). It could also be that there is a bypass of the reaction catalysed by GFPT1 existing in humans, and this bypass has diminished or eliminated the effects caused by *GFPT1* mutations. The observation of the clinical features of CMS patients with *GFPT1* mutations suggests that this bypass could only partially replace GFPT1.

Bearing in mind that not only could N-glycan antenna initiation be affected by the yield of UDP-GlcNAc, but also O-GlcNAcylation. O-GlcNAcylation is the transfer of GlcNAc to a serine or threonine residue of nuclear or cytoplasmic proteins (Hart et al., 2011). O-GlcNAcylation has been long postulated to be a unique intracellular modification (Ogawa et al., 2015; Hurtado-Guerrero et al., 2008) and the cell surface glycans seem to be more important in this study. In addition, I did not have electron transfer dissociation (ETD) MS which is an essential tool for O-GlcNAcylation analysis (Myers et al., 2013) when this project was launched, therefore I did not analyse O-GlcNAcylation in myoblasts and myotubes. However, extracellular O-GlcNAcylation has been reported (Matsuura et al., 2008). It is therefore reasonable to investigate O-GlcNAcylation in myoblast and myotubes in the future. The K_m values of the enzymes related to O-GlcNAcylation, such as intracellular O-GlcNAc transferase (OGT) and epidermal growth factor O-GlcNAc transferase (EOGT), are much smaller than those of Golgi-resident GlcNAc transferases (Ogawa et al., 2015; Haltiwanger et al., 1992; Ma et al., 2013; Lau et al., 2007). This indicates that the O-GlcNAcylation is expected to be more easily impaired by a reduction in UDP-GlcNAc concentration than N-glycans. In addition, the EOGT-catalysed O-GlcNAcylation responds to stimulation of the hexosamine biosynthesis pathway (Ogawa et al., 2015). It is likely that O-GlcNAcylation level has been affected by the *GFPT1* mutations. One function of O-GlcNAcylation in cells is to act as a sensor to regulate signalling, transcription and translation (Lau et al., 2007; Hart et al., 2011). As signalling is closely related to neuromuscular junction function it is possible that the *GFPT1* mutations might result in abnormal O-GlcNAcylation which might contribute to the human neuromuscular junction defect. This might be an explanation for the mechanism by which *GFPT1* mutations cause CMS.

The reason why O-glycans were not detected in the myotubes from *GFPT1* patients was probably that the amount of cells for analysis was limited. This explanation can be supported

by the phenomenon that fewer high mass N-glycans were observed in the *GFPT1* patient myotube samples (Figure 4.14 and Figure 4.16).

My collaborators chose other muscular diseases, namely CMS caused by *DOK7* mutations, and myopathy caused by *MTND5* mutations, Limb girdle muscular dystrophy type 2A (LGMD2A) caused by CAPN3 mutations and Pompe disease caused by *GAA* mutations, as exemplars of muscular diseases which have not been previously found to show defective N-glycosylation. However, a very recently published paper reported that Pompe disease can result in a Golgi-based glycosylation deficit in human skin fibroblast-derived induced pluripotent stem cells which were differentiated in culture to cardiomyocytes (iPSC-CMs) (Raval et al., 2015). Employing a similar glycomic strategy to ours, Raval et al. have found that there is a decreased diversity of multiantennary structures and hyposialylation in Pompe disease iPSC-CMs. However, my results did not show any obvious glycosylation defect, and the multiantennary N-glycans that I observed in my results are not evident in their control cells. This might be due to the fact that the glycosylation defects in Pompe disease are cell-specific or dependent on culture conditions.

The underlying mechanism by which *GFPT1* mutations cause CMS remains unknown. Elucidation of the mechanism could help the development of new treatment options.

Chapter 5

Glycomic profiling of trophoblasts

5. Glycomic profiling of trophoblasts

5.1. Introduction to the project

This project is in collaboration with Professor Gary F. Clark at University of Missouri and Dr. Sandra M. Blois at University Medicine of Berlin.

In order to further investigate the hu-FEDS hypothesis described in section 1.8.2 and test our hypothesis which is that functional glycan structures on human gametes (Lewis X, sialyl-Lewis X and bisecting GlcNAc) could also be detected on human trophoblasts. The N-glycomes of three trophoblast populations (CTB, STB and evCTB) were rigorously characterised. Specifically, MS analyses provided general structural information of the glycans; sialidase S digestion and MS/MS analyses were used to check whether sialyl-Lewis X structure was present; MS/MS analyses and endo-β-galactosidase digestion were employed to check whether there was Lewis X structure; linkage analysis and β1, 4-galactosyltransferase treatment were applied to verify the presence of bisecting GlcNAc.

5.2. Sample details and sample processing

The four sets of samples which were obtained from four donors were labelled as 86, 96, 114 and 117. In each set, there were one cytotrophoblast (CTB) sample and one syncytiotrophoblast (STB) sample. The detail of these samples is shown in Table 5.1.

Table 5.1 Detail of CTB and STB received from University of Missouri

Cell Type	Cell number in each sample	Number of samples
CTB86	Approximately 3 million	2
STB86	Approximately 3 million	2
CTB96	Approximately 3 million	2
STB96	Approximately 3 million	2
CTB114	Approximately 3 million	2
STB114	Approximately 3 million	2
CTB117	Approximately 3 million	2
STB117	Approximately 3 million	2

The extravillous cytotrophoblast (evCTB) samples were more authentic as they were less *in vitro* cultured and thus were closer to the real human tissues. To reduce the costs of shipping and for safety reasons, the protein extraction was carried out following the methods mentioned in section 2.2.2.1 in Germany. The detail of these evCTB samples is shown in Table 5.2.

Table 5.2 Detail of evCTB protein samples received from University Medicine Berlin

First batch			
Sample ID	**Origin**	**Cell number**	**Number of samples**
evCTB	pool of samples	12 to 15 million	1
evCTB1	pool of samples	12 to 15 million	1
evCTB3	pool of samples	12 to 15 million	1
Second batch			
Sample ID	**Origin**	**Cell number**	**Number of samples**
evCTB9W	pool of samples	Approximately 6 million	1
evCTB10W	pool of samples	Approximately 24 million	1
evCTB11W	Individual patient	Approximately 25 million	1
evCTB10 5/7	Individual patient	Approximately 18 million	1

All samples were analysed using glycomic methodologies which have been described in section 2.2: the processing of CTB and STB samples started with section 2.2.1 while the handling of evCTB samples began with section 2.2.3.

5.3.   Results

5.3.1.   CTB and STB N-glycans

5.3.1.1.   *MALDI-TOF MS analyses of the N-glycans of CTB and STB*

High quality MALDI data were obtained for the N-glycans of all the CTB and STB samples. Representative MALDI-TOF MS spectra of N-glycans of each cell type, CTB86 and STB86, are shown in Figure 5.1. The spectra show that high mannose (e.g. m/z 1580, 1784 and 1988) and complex glycans (e.g. m/z 2693 and 4226) are present in both CTB and STB. The antenna number and sialic acid numer of these complex glycans are similar to that described in section 4.3.2.1. Some common characteristics of mammalian cell N-glycomes were observed in Figure 5.1 (see section 4.3.2.1 for detailed description). In addition to the

common structures, some potentially more interesting structures, such as sialyl-Lewis X and/or A, Lewis X and/or A and bisecting GlcNAc, were observed. The sialylated N-glycans accounted for approximately 57%, 40%, 63% and 43% of the complex glycans in CTB86, CTB96, CTB114 and CTB117 respectively, and approximately 64%, 67%, 64% and 67% of the complex glycans in STB86, STB96, STB114 and STB117 respectively.
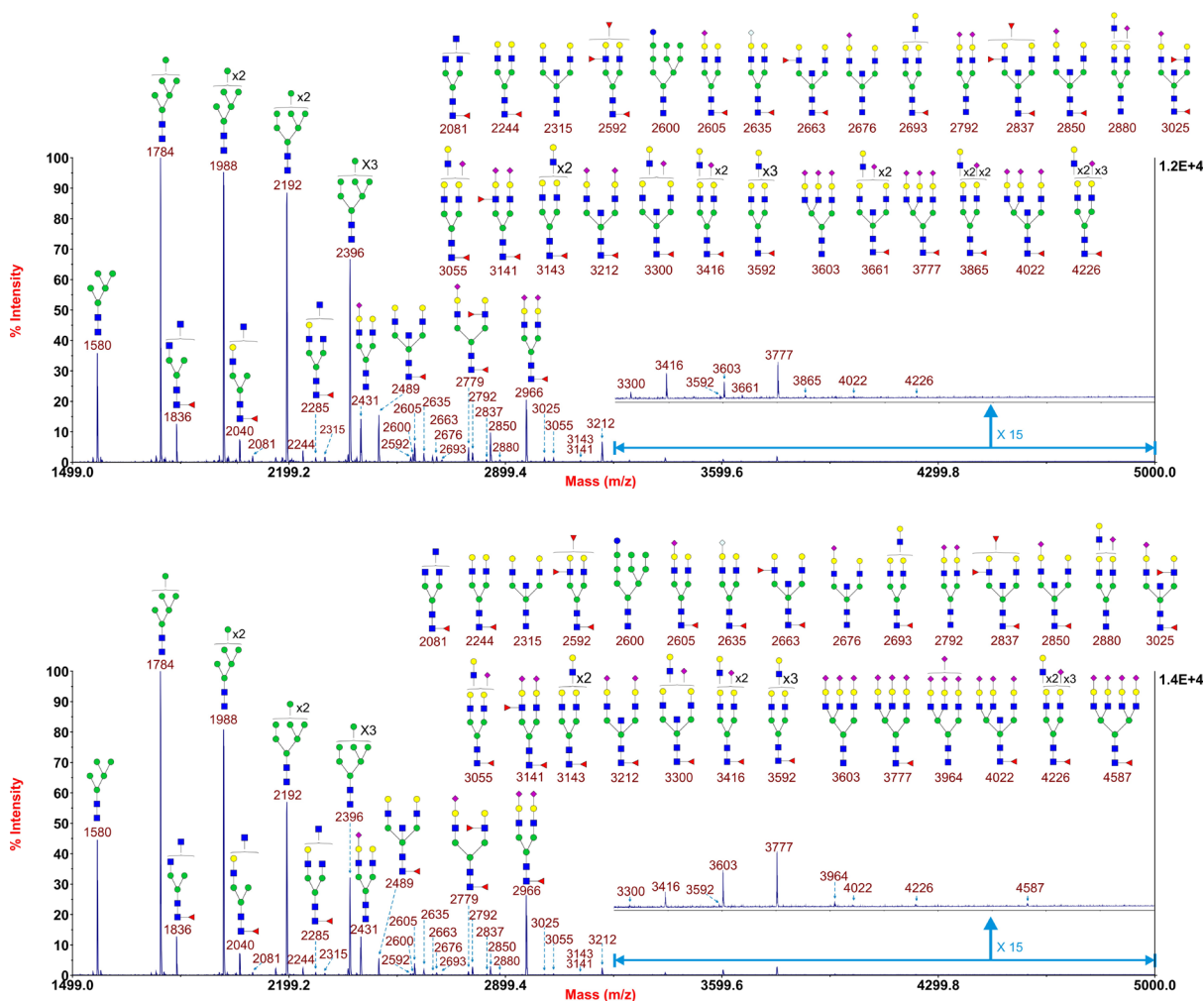
Figure 5.1 Annotated MALDI-TOF MS spectra of permethylated N-glycans from CTB86 (top) and STB86 (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. The NeuGc in the minor glycan at m/z 2635 is likely diet-derived (Tangvoranuntakul et al., 2003; Banda et al., 2012). ■ GlcNAc, ● Man, ○ Gal, ● Glc, ▲ Fuc, ◆ NeuAc, ◇ NeuGc.

The method described in section 4.3.2.1 was employed to compare fucosylation between CTB and its corresponding STB. Fucosylation level is higher in CTB when comparing the relative intensity of m/z 2779 to m/z 2605, but this difference is not observed when

comparing the relative intensity of m/z 3025 to m/z 2850 (see Table 5.3). Therefore, a firm conclusion could not be made on this fucosylation comparison.

Table 5.3 Comparisons of N-glycan fucosylation between CTB and STB

The fucosylation comparison was achieved via comparing the relative intensity (RI) of one glycan to that of the previous glycan which possesses one fewer Fuc.

Modification	RI comparison	CTB 86	STB 86	CTB 96	STB 96	CTB 114	STB 114	CTB 117	STB 117
	$RI_{(2779)}/RI_{(2605)}$	0.79	0.32	0.29	0.15	0.65	0.35	0.53	0.36
Fucosylation	$RI_{(3025)}/RI_{(2850)}$	0.13	0.12	0.18	0.17	0.09	0.09	0.09	0.11

5.3.1.2.   *MALDI-TOF/TOF MS/MS analysis of the N-glycan at m/z 2592*

MALDI-TOF/TOF MS/MS analyses were carried out to further characterize structures of peaks of interest, *i.e.* those glycans containing potential Lewis structures and sialyl-Lewis structures. Here the MS/MS data for the glycan at m/z 2592 which contains potential Lewis structures are described as an example.

As shown in Figure 5.2, the fragment ion at m/z 1955 corresponds to loss of a Lewis structure which can be either Lewis A [Galβ1,3(Fucα1,4)GlcNAc] or Lewis X [Galβ1,4(Fucα1,3)GlcNAc] or a mixture of both, and its concurrent ion is also observed at m/z 660. The signals of these two ions dominate the two spectra. These demonstrate the presence of Lewis A and/or X structure. In addition, in Figure 5.2 (bottom) the fragment ion at m/z 1781 corresponds to loss of a Lewis B [(Fucα1,2)Galβ1,3(Fucα1,4)GlcNAc] or Lewis Y [(Fucα1,2)Galβ1,4(Fucα1,3)GlcNAc] or a mixture of both, its concurrent ion is also observed at M/Z 834 though the peak signal is weak. These demonstrate the presence of Lewis B and/or Y in STB86. Fuc is confirmed to be 3-linked in the Lewis moiety as the signal at m/z 2386 corresponds to the loss of a Fuc from the C3 position of GlcNAc via beta elimination. This 3-linked Fuc is either come from the Lewis X or the Lewis Y or both. This provides evidence for the presence of potential Lewis X and/or Y. In addition, there is no detectable characteristic ion for Lewis A or B, and this can rule out the possibility that Lewis A and Lewis B are present in the CTB and STB. Therefore, these data demonstrate the presence of Lewis X in CTB and Lewis X and Y in STB.

More detailed annotation of these two spectra is shown in Figure 5.3.
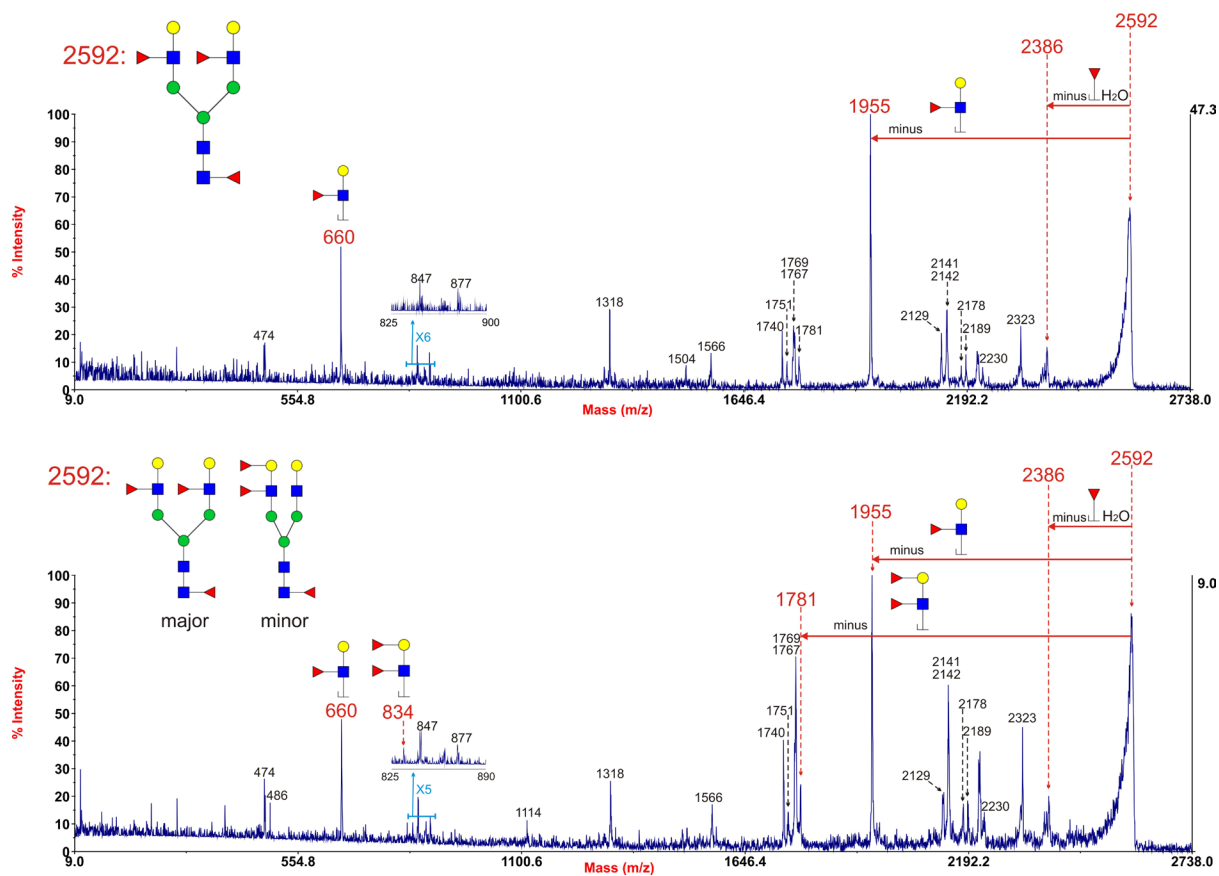
Figure 5.2 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated N-glycan at m/z 2592 in CTB86 (top) and STB86 (bottom)

Data were acquired in the form of [M+Na]+ ions. Peaks were annotated with putative fragment ions according to the molecular weight. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.
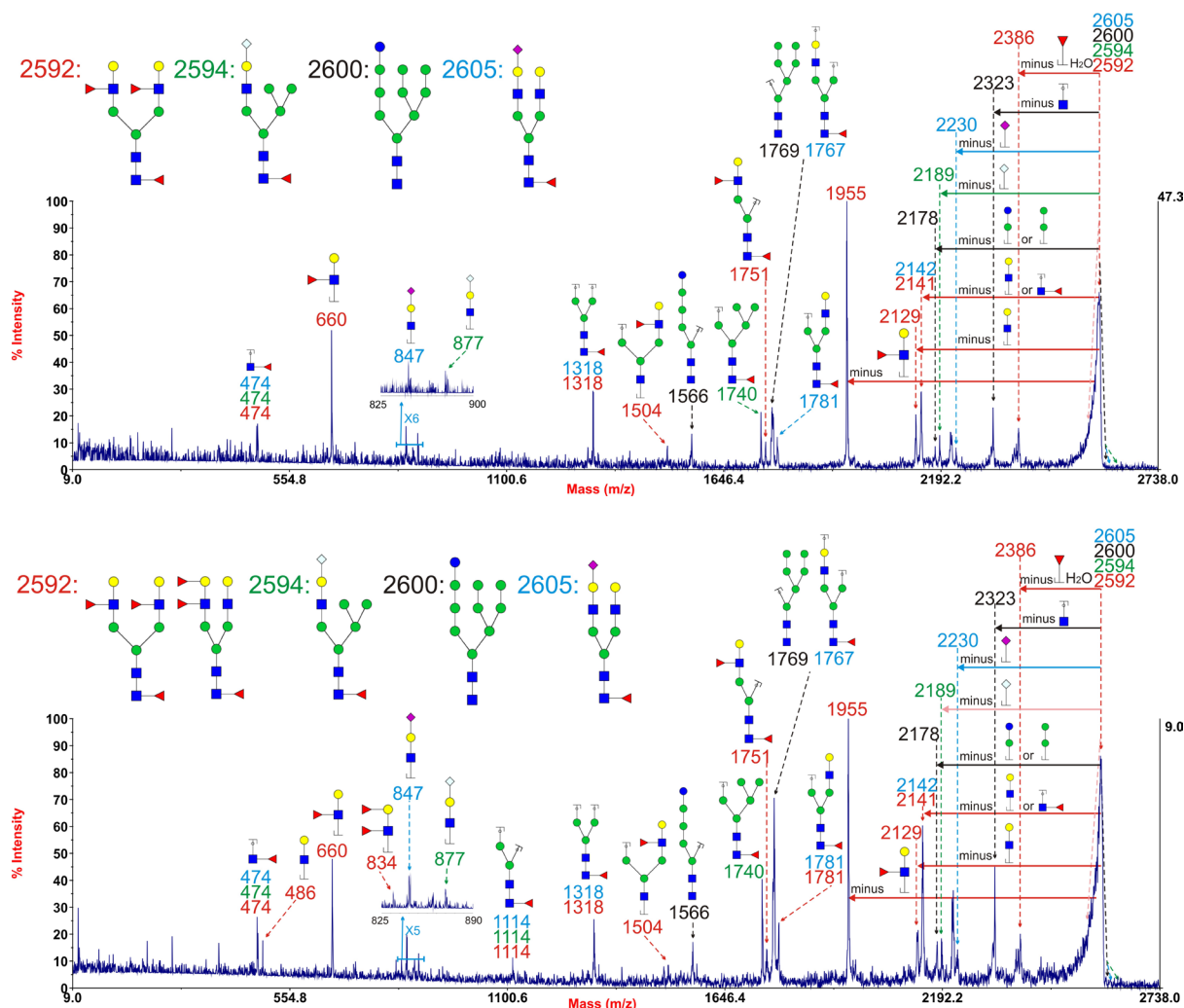
Figure 5.3 Fully annotated MALDI-TOF/TOF MS/MS spectra of permethylated N-glycan at m/z 2592 in CTB86 (top) and STB86 (bottom)

Data were acquired in the form of [M+Na]$^+$ ions. Peaks were annotated with putative fragment ions according to the molecular weight. The glycan m/z values and their corresponding fragment ions are labelled in the same colour (red, green, black or blue). ■ GlcNAc, ● Man, ○ Gal, ● Glc, ▲ Fuc, ◆ NeuAc, ◇ NeuGc.

5.3.1.3. MALDI-TOF/TOF MS/MS analysis of the N-glycan at m/z 3143

The MS/MS data for the glycan at m/z 3143 which contains potential sialyl-Lewis X and/or A are described as another example.

As observed in HeLa cells (section 3.3.1.2), the isotope peaks of the glycan at m/z 3143 span a wider mass range than expected for one single sugar composition. This implies that this peak cluster may also include LacNAc containing glycans as well as potential sialyl-Lewis X

179

and/or A containing glycans which differ by 2 Da. MS/MS analyses were carried out to check this.

The spectra shown in Figure 5.4 are similar to that in Figure 3.5 (top), see section 3.3.1.2 for detailed explanation. This figure provides evidence for the presence of potential sialyl-Lewis X. A comparison of the relative intensity of characteristic fragment ions (m/z 847, 1021, 2142, 2316, 2765 and 2934) of potential sialyl-Lewis X containing glycan to the sum of all fragment ions shows that the potential sialyl-Lewis X containing glycan accounts for approximately 60% and 56% of the glycan mixture in CTB86 and STB86 respectively. Similarly, a comparison of the relative intensity of the characteristic fragment ions (m/z 935, 1781 and 2230) to the sum of all fragment ions shows that the polyLacNAc containing glycan accounts for approximately 7% of the glycan mixture in both CTB86 and STB86. The remaining should be tetraantennary glycan with each antenna possessing one LacNAc.

To further check the presence of sialyl-Lewis X, it is necessary to determine the linkage of NeuAc, and this can be achieved via sialidase S digestion.

Figure 5.4 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated N-glycan peak centred at m/z 3142 in the CTB86 (top) and STB86 (bottom)

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

5.3.1.4. Sialidase S digestion of the N-glycans of CTB and STB

Digestion of both CTB and STB with sialidase S resulted in partial desialylation, with only a handful of minor sialylated N-glycans at m/z 2431, 2605 and 2850 being observed (Figure 5.5). The glycan at m/z 2489 is the major component in the complex N-glycans. Comparison of this figure and Figure 5.1 indicated that both CTB and STB mainly contain α2,3 linked

NeuAc, which could be observed via a significant but not complete loss of NeuAc residues from the assigned structures. After the digestion, the isotope peak cluster around m/z 3142 in both CTB and STB was consistent with a single composition, which was the same as described in section 3.3.1.4. In addition, some high mass glycans, such as m/z 4041, 4490 and 4939, were not detected previously but observed after the enzymatic digestion. Their putative structures suggest the existence of α2,3 sialylated polyLacNAc in both CTB and STB. The desialylated high mass components were checked by MS/MS analyses.

Figure 5.5 Annotated MALDI-TOF MS spectra of permethylated sialidase S treated N-glycans from CTB86 (top) and STB86 (bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. See legend to Figure 5.1 for explanation of the NeuGc containing glycan. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc, ◇ NeuGc.

As shown in Figure 5.6, the fragment ions corresponding to loss of LacNAc, diLacNAc and triLacNAc were observed at m/z 2679, 2230 and 1781 respectively, while fragment ions corresponding to loss of NeuAc or NeuAc related structure were not detected. The fragment ion at m/z 2691 corresponds to the loss of a fucosylated GlcNAc. This fragment ion has been previously observed in Figure 5.4. There is no other fragment ion corresponding to Fuc or Fuc containing residues. This suggested that the NeuAc in the previous potential sialyl-Lewis X glycan (m/z 3141) had been cleaved and thus the NeuAc should be α2,3 linked, after sialidase S digestion the m/z valude of the glycan has shifted. This demonstrates the presence of sialyl-Lewis X in both CTB and STB.

Figure 5.6 Annotated MALDI-TOF/TOF MS/MS spectra of permethylated sialidase S treated N-glycan at m/z 3143 in CTB86 (top) and STB86 (bottom)

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

5.3.1.6. ***MALDI-TOF/TOF MS/MS analysis of the N-glycan at m/z 4939 after sialidase S digestion***

As mentioned previously, high mass glycans, such as such as m/z 4041, 4490 and 4939, may contain polyLacNAc. Therefore, MS/MS analyses were carried out on these glycans. As an example, data obtained from the sialidase S digested glycan at m/z 4939 are shown in Figure 5.7. The fragment ions corresponding to loss of LacNAc, diLacNAc, triLacNAc, tetraLacNAc and pentaLacNAc were observed at m/z 4476, 4027, 3578, 3129 and 2679 respectively. These fragment ions dominate the spectra. Some of their concurrent ions were also observed at m/z 486, 935, 1385 and 1834. These clearly demonstrate the presence of polyLacNAc. The maximum number of the LacNAc repeating unit observed in the CTB and STB MS/MS data is six (see Figure 5.7 bottom), although the signal is weak.
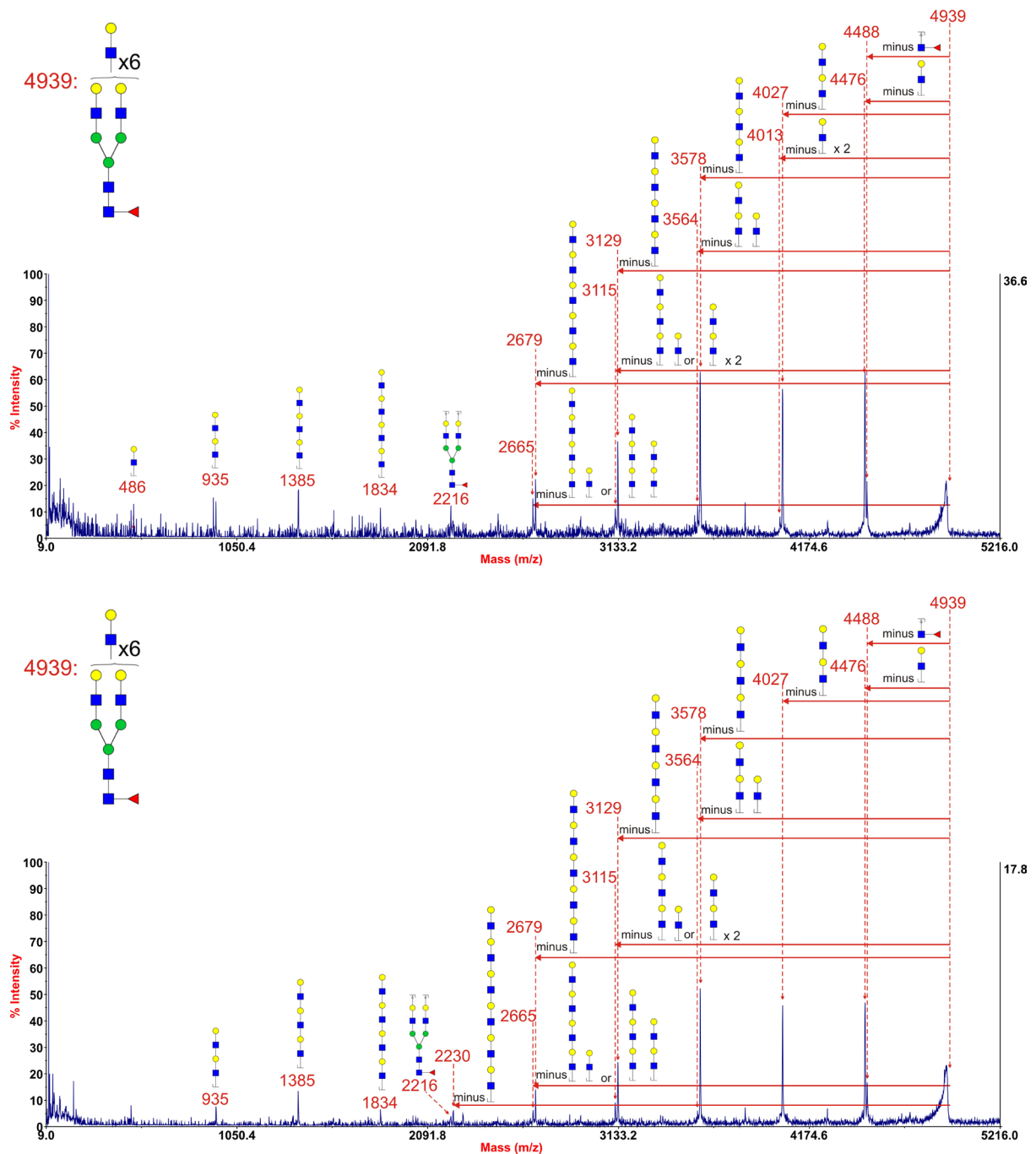
Figure 5.7 Annotated MALDI-TOF MS/MS spectra of permethylated sialidase S digested N-glycan at 4939 in CTB114 (top) and STB114 (bottom)

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

5.3.1.7.   *Linkage analysis of the N-glycans of CTB and STB*

In addition to Lewis X, Lewis Y, sialyl-Lewis X and polyLacNAc, bisecting GlcNAc may also exist in the CTB and STB. However, MS/MS analysis is not able to check the presence

of bisecting GlcNAc as its fragment ion cannot be distinguished from that of a truncated antenna.

As shown in Figure 1.6, bisecting GlcNAc is linked to 3, 4, 6-linked Man, therefore the presence of 3, 4, 6-linked Man can reflect the existence of bisecting GlcNAc. GC-MS is able to check the presence of 3, 4, 6-linked Man. Table 5.4 summarizes the result of GC-MS linkage analysis of the glycan derivatives from CTB117 and STB117 and demonstrates the presence of 3, 4, 6-linked Man, and this provides evidence for the existence of bisecting GlcNAc.

Table 5.4 Summary of GC-MS linkage analysis of partially methylated alditol acetates derived from the 50% acetonitrile fraction of permethylated N-glycans of CTB117 and STB117

The elution time is indicated in minutes and the relative abundance is normalised to the abundance of 2-linked mannose (major component) which is designated as 1.

Elution time, min (CTB117)	Elution time, min (STB117)	Characteristic fragment ions	Assignments	Relative abundance (CTB117)	Relative abundance (STB117)
16.95	16.90	102, 115, 118, 131, 162, 175	Terminal Fuc	0.16	0.14
18.45	18.40	102, 118, 129, 145, 161, 205	Terminal Man	0.68	0.62
18.71	18.67	102, 118, 129, 145, 161, 205	Terminal Gal	0.15	0.17
19.62	19.56	129, 130, 161, 190, 234	2-linked Man	1	1
19.90	19.85	118, 129, 161, 203, 234	3-linked Gal	0.07	0.10
21.18	21.14	87, 88, 129, 130, 189, 190	2,6-linked Man	0.05	0.05
21.34	21.30	118, 129, 189, 202, 234	3,6-linked Man	0.33	0.34
21.80	21.76	118, 139, 259, 333	3,4,6-linked Man	0.08	0.07
22.27	22.23	117, 129, 145, 205, 247	Terminal GlcNAc	0.04	0.04
23.15	23.12	117, 159, 233	4-linked GlcNAc	0.22	0.39
24.00	23.96	117, 159, 346	3,4-linked GlcNAc	0.03	0.03
24.46	24.42	117, 159, 261	4,6-linked GlcNAc	0.04	0.08

5.3.1.8.   The N-glycans of CTB and STB treated with β1, 4-galactosyltransferase

In addition to GC-MS linkage analysis, β1,4-galactosyltransferase incubation was also carried out to prove the presence of bisecting GlcNAc. β1,4-galactosyltransferase can transfer a Gal from UDP-Gal to a GlcNAc producing the disaccharide unit, Galβ1,4GlcNAc.

However, if the GlcNAc is bisecting, it will not be modified by this enzyme (Qasba et al., 2008; Narasimhan et al., 1985). This incubation was carried out on the same samples (117C and 117S) which had been investigated by linkage analysis. Glycans at m/z 2489, 2850 and 3212 were chosen as the research subjects in this experiment as all of them have a GlcNAc which could be either truncated or bisecting. Figure 5.8 shows that after β1,4-galactosyltransferase incubation, the relative intensities of glycans at m/z 2489, 2850 and 3212 in both CTB and STB did not alter significantly compared to that of glycans at m/z 2693, 3055 and 3416 respectively. This indicates that β1,4-galactosyltransferase could not modify the GlcNAc in these glycans and thus the GlcNAc present in these glycans are bisecting.
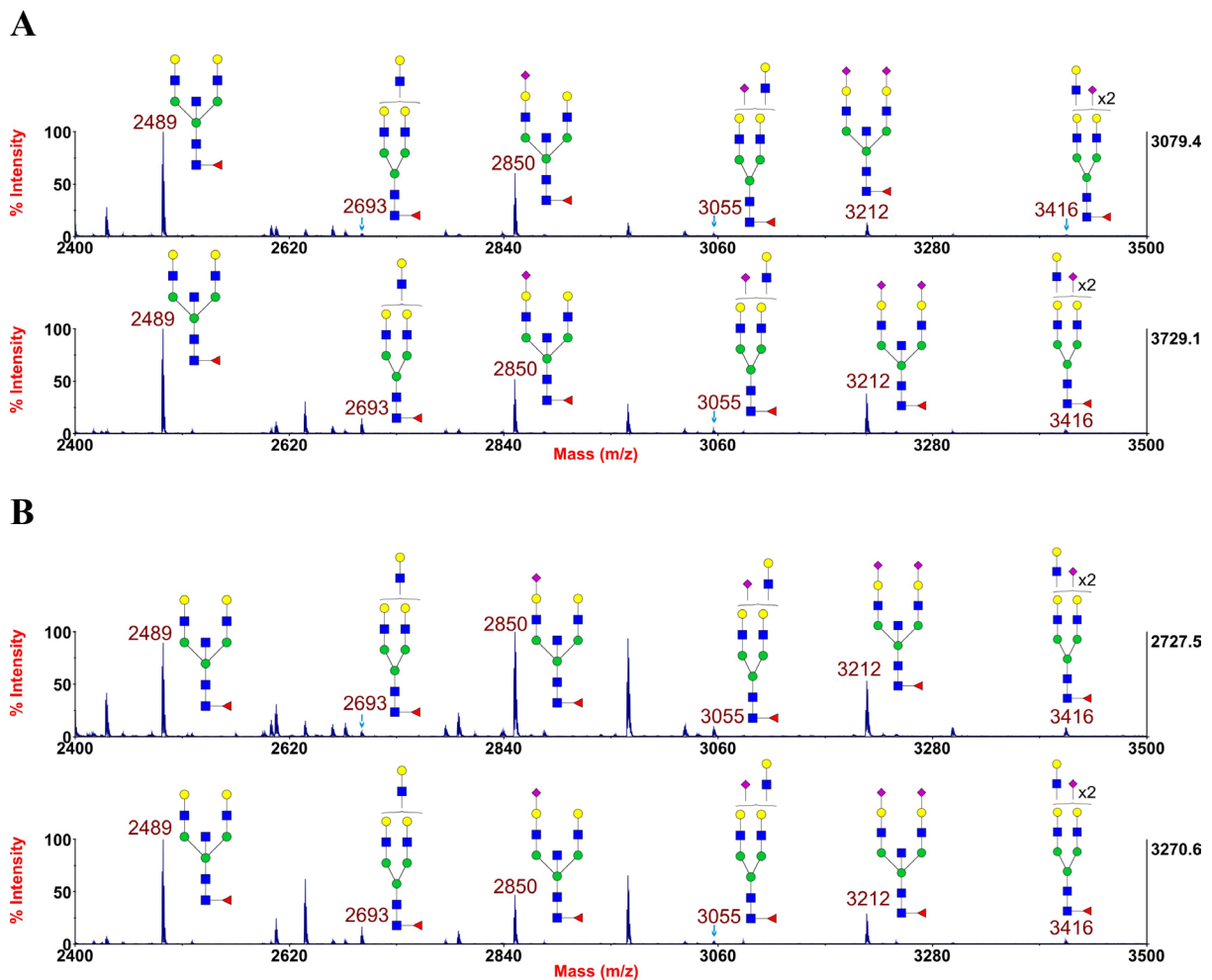
Figure 5.8 Annotated MALDI-TOF MS spectra of permethylated N-glycans from CTB117 (A, top) and STB117 (B, top) and permethylated β1,4-galactosyltransferase incubated N-glycans from CTB117 (A, bottom) and STB117 (B, bottom)

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

The bisecting GlcNAc-containing N-glycans account for approximately 32%, 15%, 50% and 59% of the complex glycans in CTB86, CTB96, CTB114 and CTB117 respectively, and approximately 15%, 22%, 41% and 49% of the complex glycans in STB86, STB96, STB114 and STB117 respectively.

5.3.2.   evCTB N-glycans

5.3.2.1.   MALDI-TOF MS analysis of the N-glycans of evCTB

There were seven evCTB samples in total. They were in turn labelled as evCTB, evCTB1, evCTB3, evCTB9W, evCTB10W, evCTB11W and evCTB10 5/7. All the samples were analysed using the same method described in section 2.2. High quality MALDI data were obtained for the N-glycans of evCTB. The MALDI-TOF MS spectrum of N-glycans of one representative sample, evCTB10 5/7 is shown in Figure 5.9. The spectrum shows that high mannose and complex glycans are present in evCTB. Common characteristics of mammalian cell N-glycomes and potentially more interesting structures were observed in Figure 5.9 (see section 5.3.1.1 for detailed description). The sialylated N-glycans accounted for approximately 65%, 68%, 40%, 48%, 14%, 12% and 72% of the complex glycans in evCTB, evCTB1, evCTB3, evCTB9W, evCTB10W, evCTB11W and evCTB10 5/7 respectively.
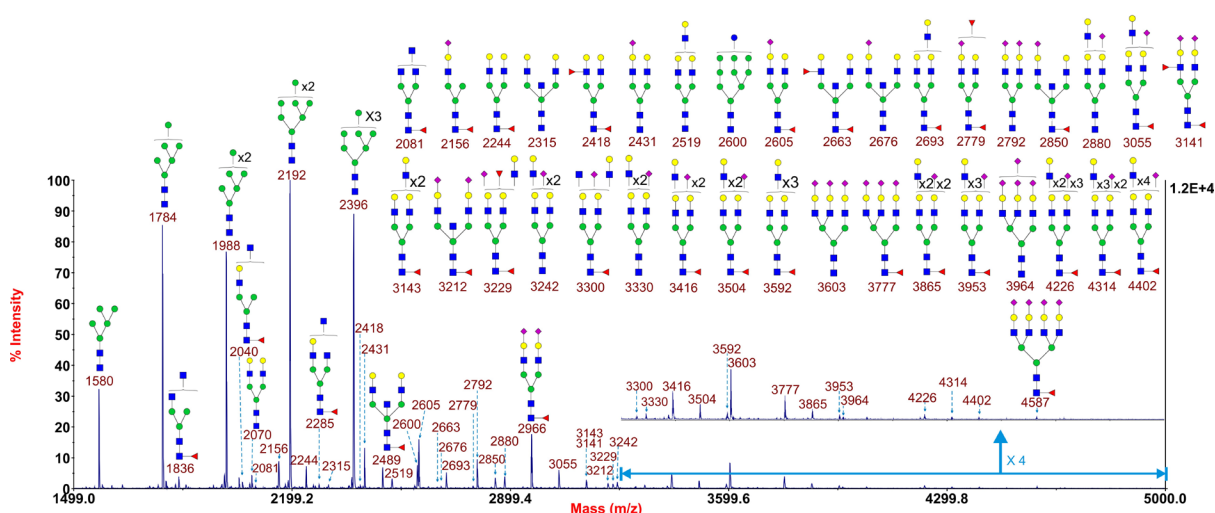
Figure 5.9 Annotated MALDI-TOF MS spectrum of permethylated N-glycans from evCTB10 5/7

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are $[M+Na]^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ● Glc, ▲ Fuc, ◆ NeuAc.

There was no significant alteration when comparing the fucosylation in evCTB (Table 5.5) to that in CTB and STB (Table 5.3).

Table 5.5 Comparisons of N-glycan fucosylation

See legend to Table 5.3 for explanation of the comparison. NA, not available, which is due to the fact that the RI of the glycan is too low to be used in the calculation.

Modification	RI comparison	evCTB	evCTB1	evCTB3	evCTB 9W	evCTB 10W	evCTB 11W	evCTB 10 5/7
Fucosylation	$RI_{(2779)}/RI_{(2605)}$	1.10	0.57	0.45	0.25	0.68	0.56	0.82
	$RI_{(3025)}/RI_{(2850)}$	0.28	0.14	0.38	0.49	NA	NA	0.06

5.3.2.2. *MALDI-TOF/TOF MS/MS analysis of the N-glycan at m/z 3143*

In evCTB the isotope peaks of the glycan at m/z 3143 also span a wider than expected, which is the same as the observation in section 3.3.1.2, therefore MS/MS analysis was carried out.

The MS/MS spectrum shown in Figure 5.10 is similar to that in Figure 3.5 (top), see section 3.3.1.2 for detailed explanation. The following figure provides evidence for the presence of potential sialyl-Lewis X. Following the calculation method described in section 5.3.1.3, the potential sialyl-Lewis X containing glycan accounts for approximately 15% of the glycan

189

mixture, the poly-LacNAc containing glycan also accounts for approximately 15%. Compared to Figure 5.4 (CTB and STB), the relative abundance of sialyl-Lewis X to polylacNAc is decreased in evCTB (approximately 9.22 in CTB, 8.33 in STB, 1.00 in evCTB).
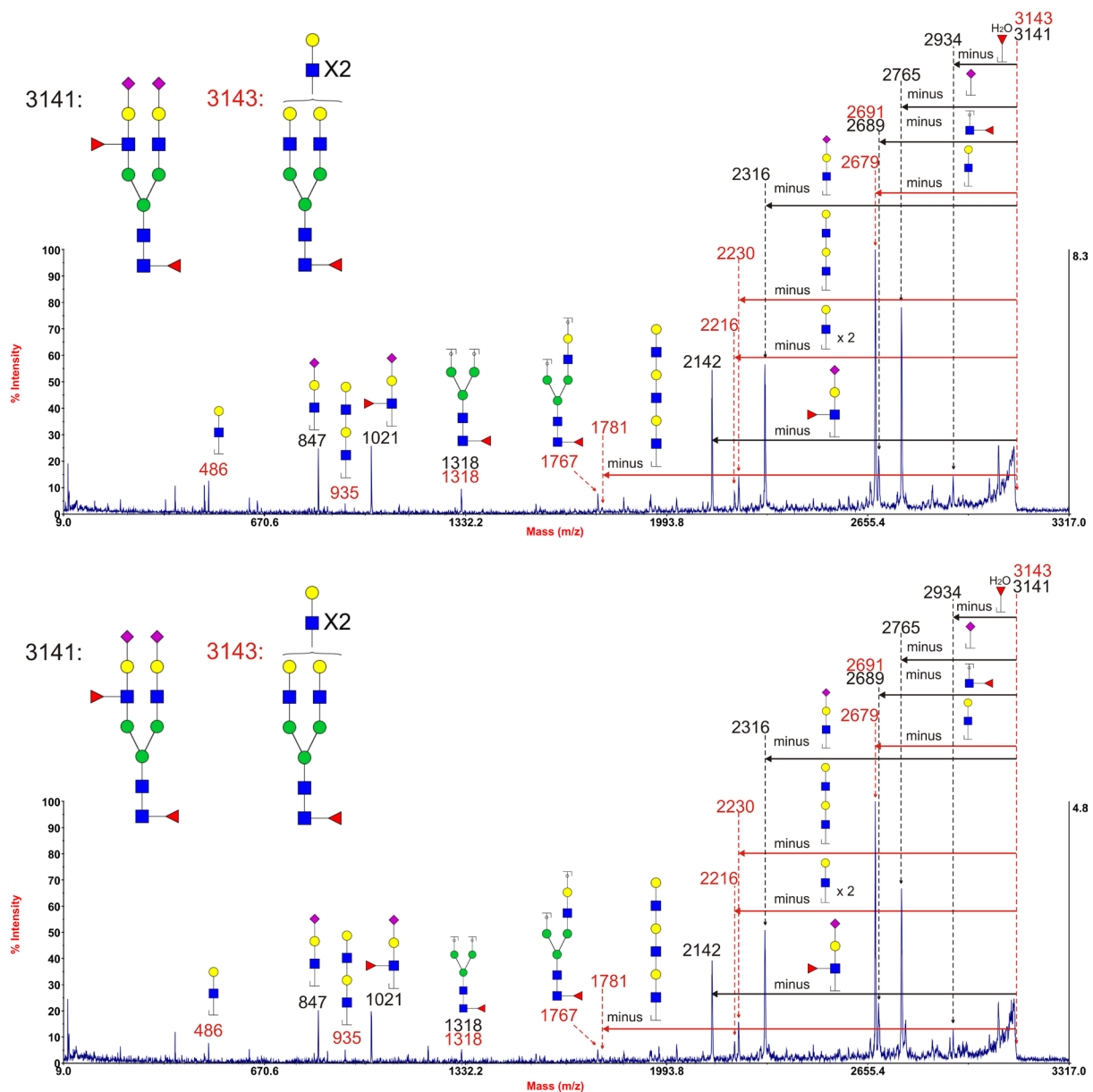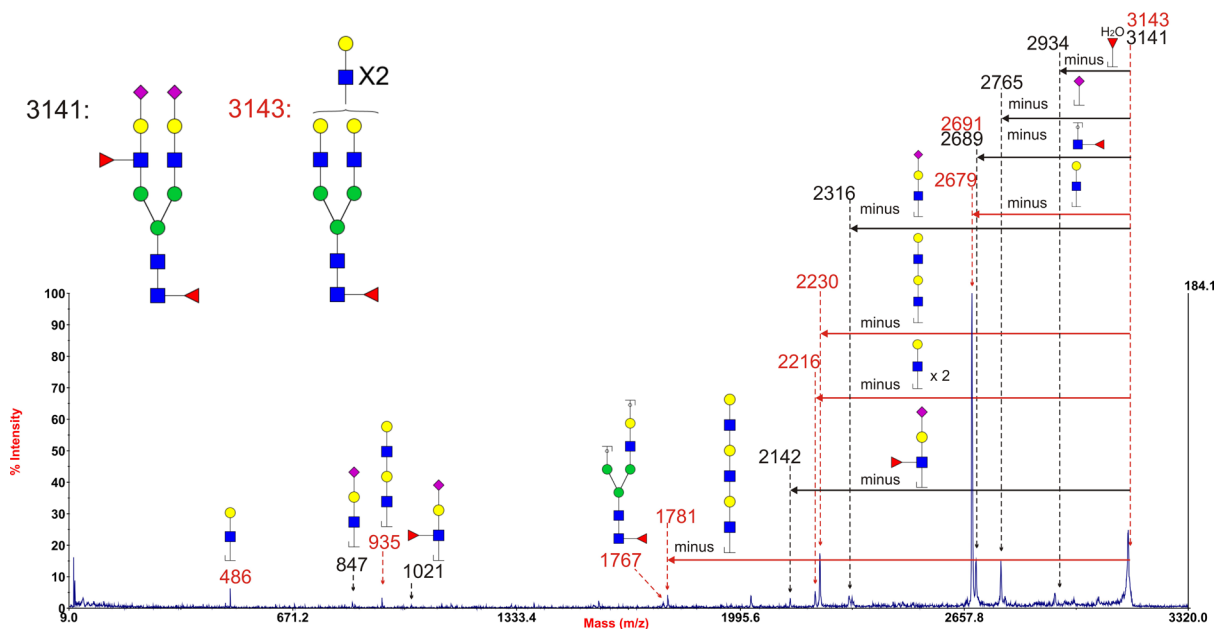

Figure 5.10 Annotated MALDI-TOF/TOF MS/MS spectrum of permethylated N-glycan peak centred at m/z 3142 in the evCTB10 5/7

Data were acquired in the form of $[M+Na]^+$ ions. See legend to Figure 3.5 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

5.3.2.3.   *Sialidase S digestion of the N-glycans of evCTB*

The MALDI spectrum of sialidase S digested evCTB N-glycans is shown in Figure 5.11. Comparison of this figure with Figure 5.9 indicated that the all NeuAcs in N-glycans of evCTB were α2,3 linked as a complete loss of NeuAc was observed. Unlike the spectra in Figure 5.5 (sialidase S digested N-glycans of CTB and STB) which are dominated by high mannose and biantennary glycans, the following spectrum is dominated by glycans at m/z 2693 and 3143.

As observed in CTB and STB (section 5.3.1.4), the high mass glycans, such as m/z 4041, 4490 and 4939, were not detected previously (in Figure 5.5) but observed after the enzymatic digestion in evCTB. This suggests the existence of polyLacNAc, which was checked by MS/MS analyses.
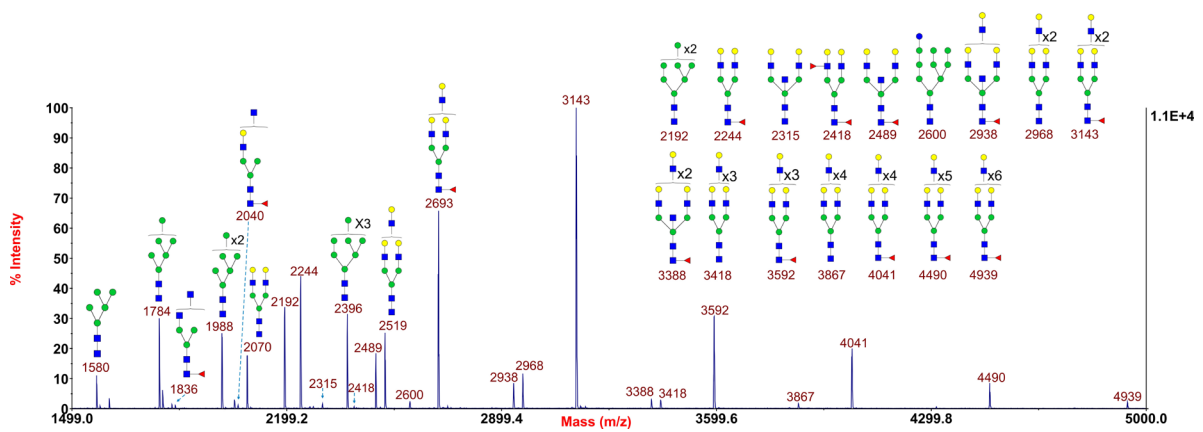
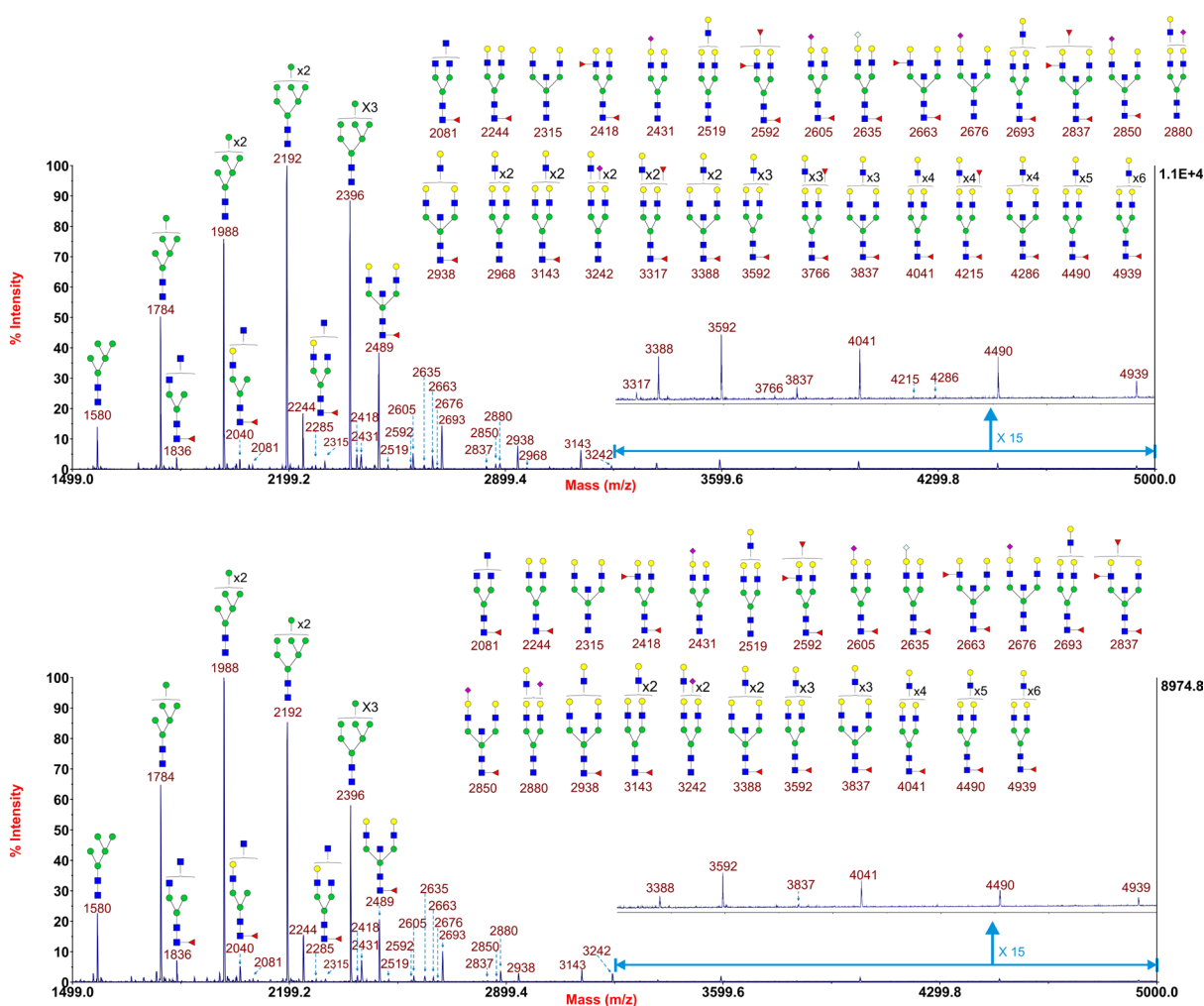**Figure 5.11 Annotated MALDI-TOF MS spectrum of permethylated sialidase S treated N-glycans from evCTB10 5/7**

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

5.3.2.4. MALDI-TOF/TOF MS/MS analysis of the N-glycan at m/z 3143 after sialidase S digestion

After the digestion, the isotope peak cluster around m/z 3142 in evCTB was consistent with a single composition, and this was consistent with the description in section 3.3.1.4 in HeLa cells. The MS/MS spectrum shown in Figure 5.12 is similar to that in Figure 3.8, see section 3.3.1.4 for detailed explanation.

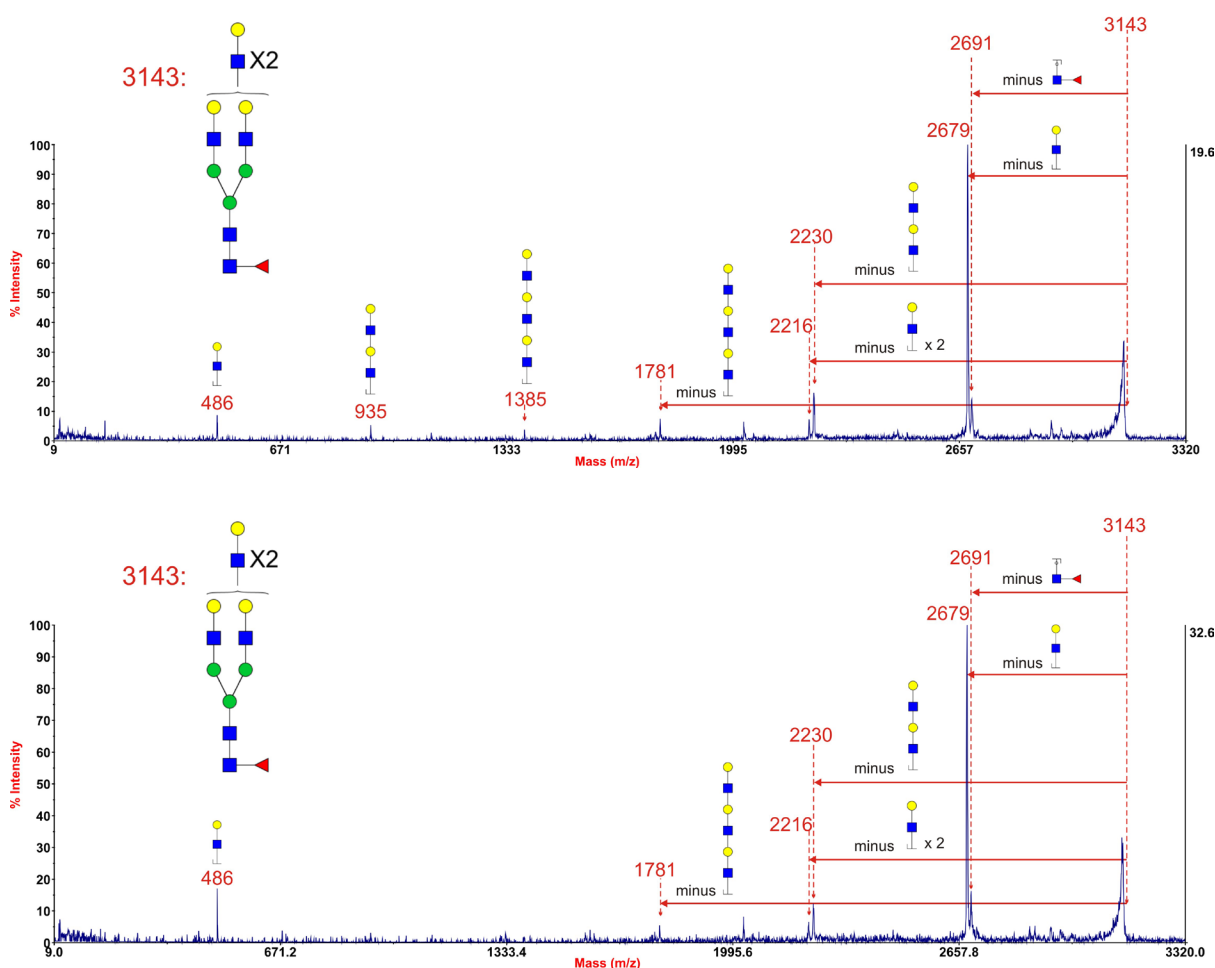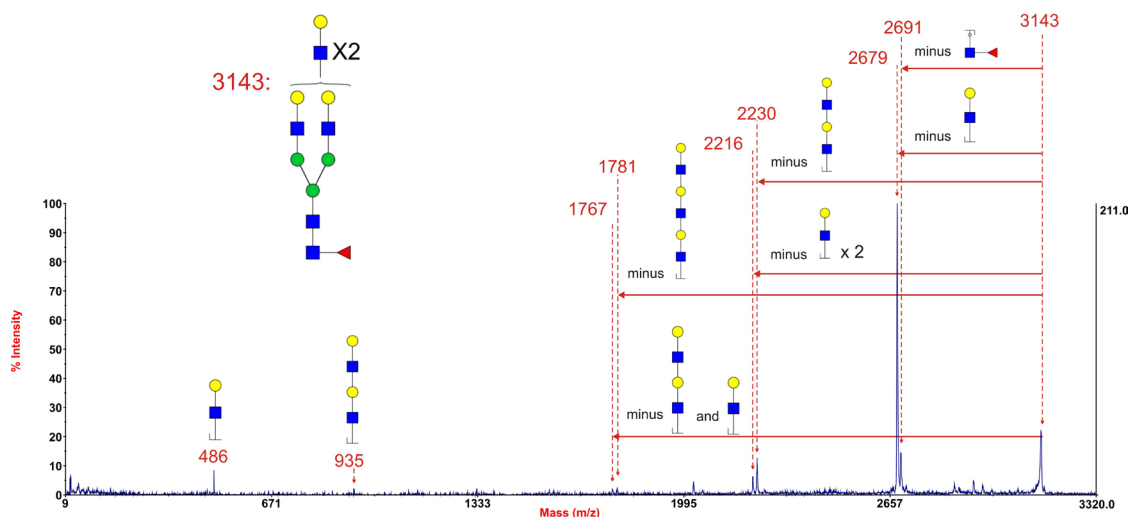**Figure 5.12 Annotated MALDI-TOF/TOF MS/MS spectrum of permethylated sialidase S treated N-glycan at m/z 3143 in evCTB10 5/7**

Data were acquired in the form of [M+Na]$^+$ ions. See legend to Figure 3.5 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

The MS/MS spectrum shown in Figure 5.13 is similar to that in Figure 5.7; detailed explanation could be found in section 5.3.1.6. This figure clearly provides evidence for the presence of polyLacNAc in evCTB N-glycans. The maximum number of the LacNAc unit in a polyLacNAc observed here is four.
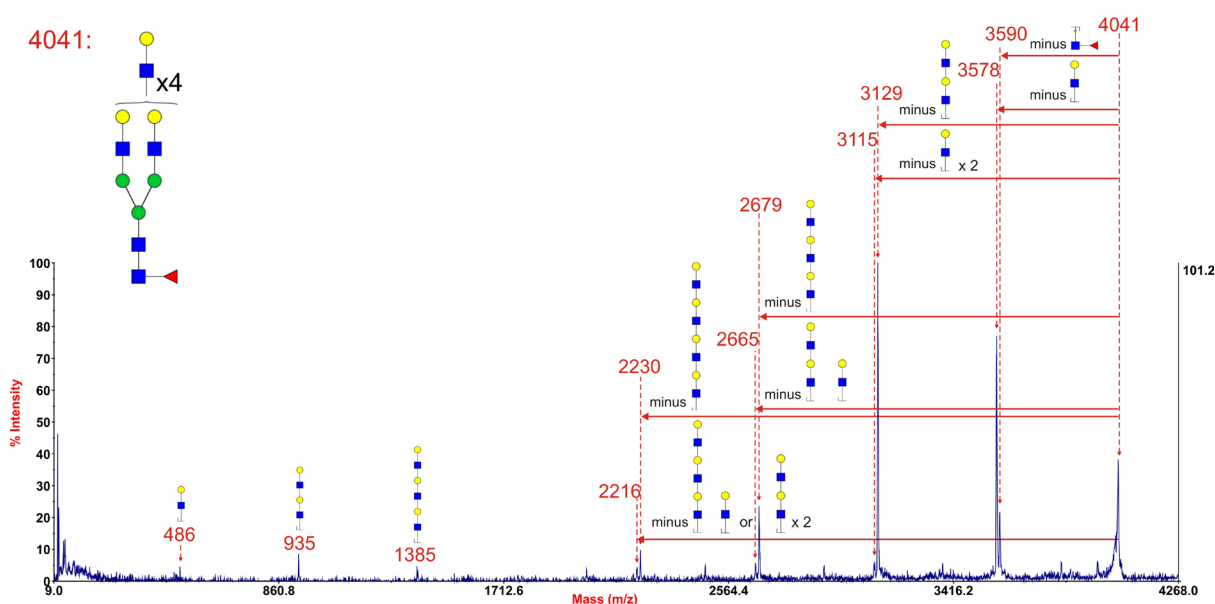
Figure 5.13 Annotated MALDI-TOF MS spectrum of permethylated sialidase S digested N-glycan at 4041 in evCTB 10 5/7

Data were acquired in the form of $[M+Na]^+$ ions. See legend to Figure 3.5 for explanation of assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc.

### 5.3.2.6.	*The N-glycans of evCTB and STB treated by endo-β-galactosidase*

Whether polyLacNAc is present in the evCTB can also be verified using endo-β-galactosidase which is an enzyme hydrolysing internal β1,4 galactosidic linkage in this favored repeating unit $[GlcNAc\beta1,3Gal\beta1,4]_n$ (Scudder et al., 1983). In addition, digestion of glycans with this enzyme can reveal the relative abundances of the non-reducing terminal structures to the core structures. N-glycans from sample evCTB11W were treated with this enzyme; some of the small resulting glycans were washed out by 35% acetonitrile, the MS

spectrum of which is shown in Figure 5.15 (top). A comparison of Figure 5.14 with Figure 5.15 shows that endo-β-galactosidase treatment gives rise to signals at m/z 518, 722, 896 and 1084. These structures could be the middle part of an antenna (m/z 518) Gal-terminated terminal (m/z 722), sialylated terminal (m/z 1084) or fucosylated terminal (m/z 896). These data provided firm evidence for the presence of polyLacNAc capped with NeuAc or Fuc, however, the terminal structure sialyl-Lewis X was not detected.
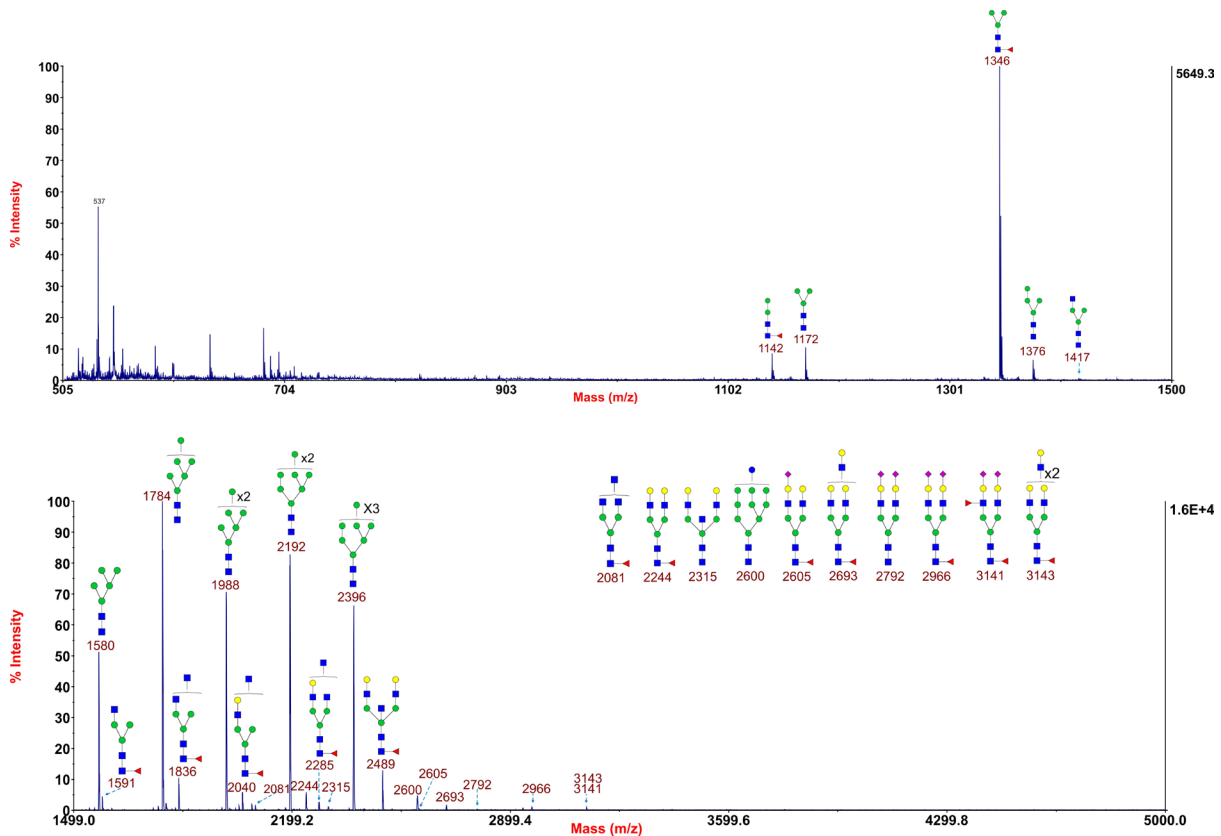
Figure 5.14 Annotated MALDI-TOF MS spectra of permethylated N-glycans from evCTB11W

Profiles were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column. All ions are [M+Na]$^{+}$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.
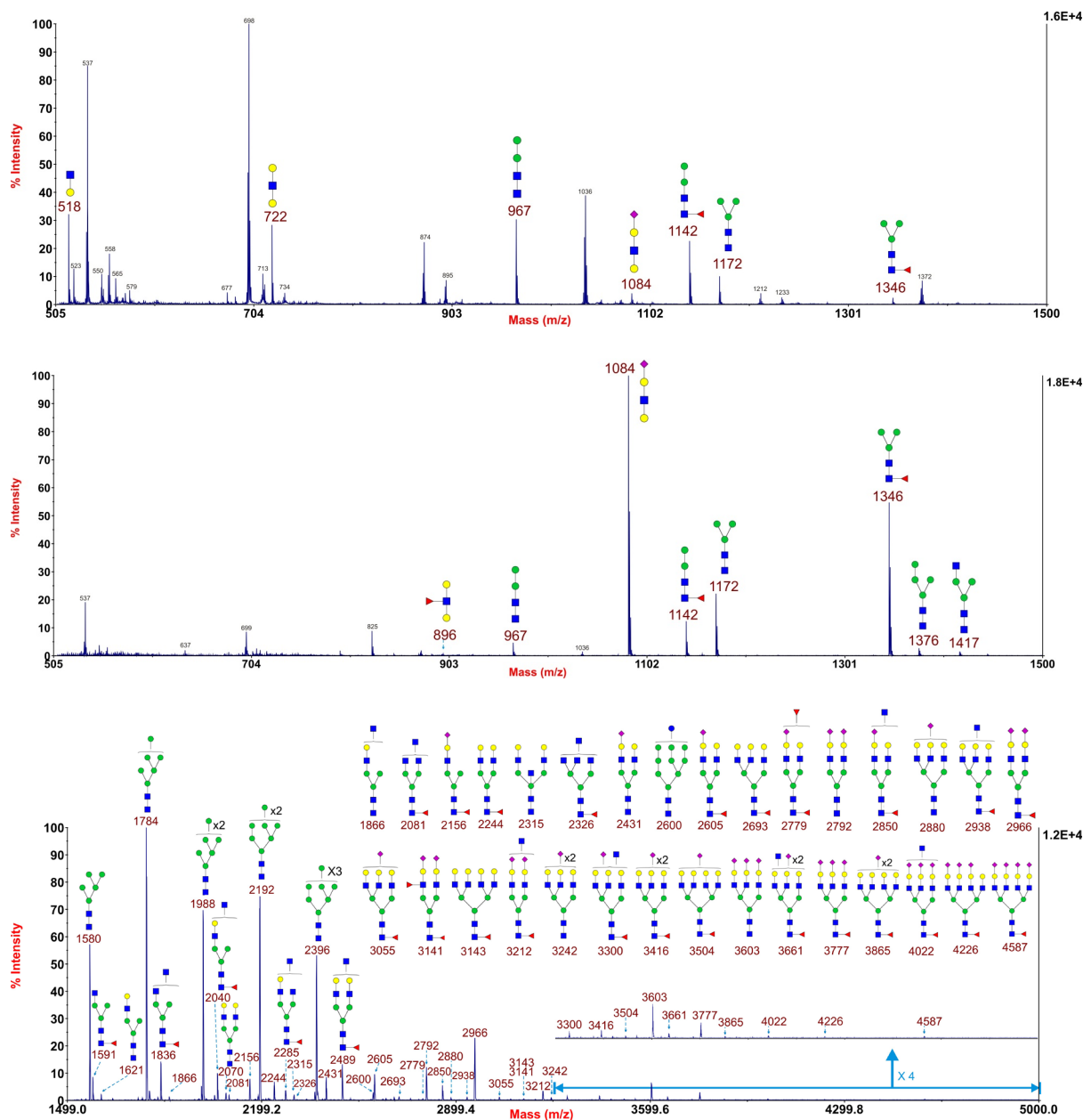
Figure 5.15 Annotated MALDI-TOF MS spectra of permethylated endo-β-galactosidase treated N-glycans from evCTB11W

Glycan profile shown in the top panel was obtained from the 35% acetonitrile fraction from a C18 Sep-Pak column, profiles shown in middle and bottom panels were obtained from the 50% acetonitrile fraction from a C18 Sep-Pak column, All ions are [M+Na]$^+$. See legend to Figure 3.2 for explanation of structure assignments. ■ GlcNAc, ● Man, ○ Gal, ▲ Fuc, ◆ NeuAc.

5.3.2.7. *MALDI-TOF/TOF MS/MS analysis of the N-glycan at m/z 896 after endo-β-galactosidase digestion*

The signal of the glycan at m/z 896 in Figure 5.15 (middle) was very weak and it was only observed in Figure 5.15 (middle), so it was necessary to check whether it was a glycan.

Therefore MS/MS analysis was carried out. As shown in Figure 5.16, the fragment ion at m/z 660 corresponds to loss of a Gal, and its concurrent ion at m/z 259 was observed. The ion at m/z 660 corresponds to the B ion of Lewis structure. In addition, the signal at m/z 690 corresponds to the loss of a Fuc via elimination. All these provide evidence for the existence of the Lewis X structure.
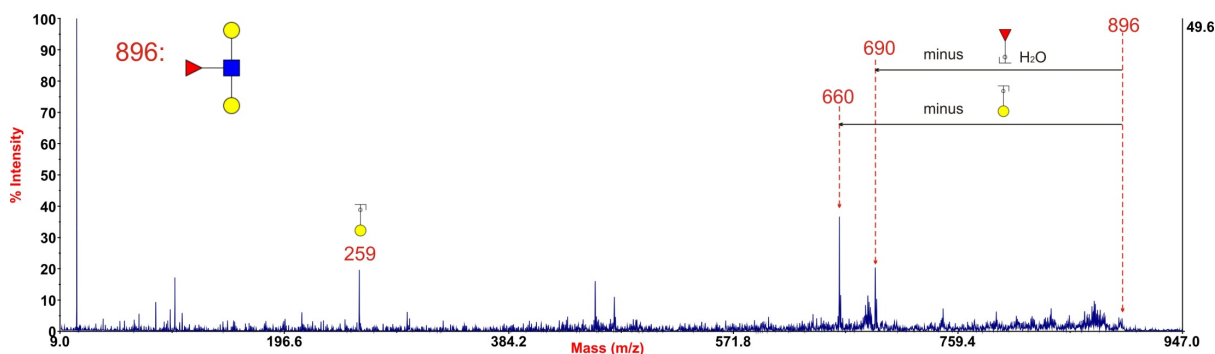
Figure 5.16 Annotated MALDI-TOF/TOF MS/MS spectrum of permethylated N-glycan at m/z 896 from evCTB11W

Data were acquired in the form of [M+Na]$^+$ ions. Peaks were annotated with putative fragment ions according to the molecular weight. ■ GlcNAc, ● Gal, ▲ Fuc.

5.3.3.   CTB, STB and evCTB

The following table summarizes the difference between CTB, STB and evCTB.

Table 5.6 A summary of the difference between CTB, STB and evCTB

Cell type	CTB	STB	evCTB
Locus	Foetal part	Foetal part	Maternal part
Sialic acid linkage	Mainly α2,3 linked	Mainly α2,3 linked	all α2,3 linked
Ratio of sialylated N-glycans in complex glycans	51%	65%	59%
Relative abundance of sialyl-Lewis X to polyLacNAc	Higher than evCTB	Higher than evCTB	Lower than CTB and STB
Ratio of glycans with at least three LacNAc units (could be potential polyLacNAc)	Lower than evCTB	Lower than evCTB	Higher than CTB and STB
Ratio of bisecting GlcNAc-containing N-glycans	39%	32%	19%

5.4.   Discussion

The human placenta is an interface between mother and foetus. It is responsible for not only nutrient exchange but also protecting the foetus from being injured by maternal immune cells. However, the mechanism underlying the protection is still not clear. Knowing the glycans expressed by the CTB, STB and evCTB could help us to better understand how glycans might be involved in this protection.

Useful information was obtained from MALDI-TOF, MALDI-TOF/TOF, GC-MS linkage analysis, sialidase S digestion, endo-β-galactosidase digestion, β1,4-galactosyltransferase incubation experiments and the knowledge of N-glycan biosynthesis pathway. Although sample quantity limitations prevented the observation of very high mass glycans, an abundance of good quality data were still obtained. These data provide evidence for the presence of α2,3-linked NeuAc, Lewis X, sialyl-Lewis X, polyLacNAc and bisecting GlcNAc in CTB, STB and evCTB.

CTB and STB were obtained from normal term placenta (approximately 39-40 weeks) and evCTB were isolated from 10-12 weeks old placenta.

In CTB, STB and evCTB, the NeuAcs in N-glycans are mainly α2,3 linked (Figure 5.1, Figure 5.5, Figure 5.9 and Figure 5.11). The sialylated N-glycans accounted for approximately 51% (mean value of four samples) of the complex glycans in CTB and approximately 65% (mean value of four samples) in STB. It has been observed that two of the evCTB samples have virtually no sialic acid (evCTB10W and 11W), and this could be caused by the loss of NeuAc during sample handling prior to sending to our laboratory. Excluding these two samples, the sialylated N-glycans accounted for approximately 59% (mean value of five samples) of the complex glycans in evCTB. This shows that the sialylation has not significantly changed among these three types of trophoblasts. More than half of the N-glycans in these trophoblasts are sialylated indicates the important role of sialylation.

Lewis structures (Lewis and sialyl-Lewis) were present in CTB, STB and evCTB, and this is supported by Figure 5.2, Figure 5.4, Figure 5.6 from CTB and STB and Figure 5.10, Figure 5.12, Figure 5.16 from evCTB. A comparsion of Figure 5.4 and Figure 5.10 showed that the relative abundance of sialyl-Lewis X (from the glycan at m/z 3141) to polyLacNAc (from the

glycan at m/z 3143) was higher in CTB and STB; this implied that both of sialyl-Lewis X and polyLacNAc were essential for these trophoblasts. The relative abundance variation in CTB, STB and evCTB suggested their different glycan-related functions.

PolyLacNAc antennae were observed in CTB, STB (Figure 5.7) and evCTB (Figure 5.13). A comparison of MALDI MS spectra of sialidase S digested N-glycans of CTB and STB (Figure 5.5) to that of evCTB (Figure 5.11) showed that evCTB possessed higher levels of multiantennary and/or polyLacNAc N-glycans.

In CTB and STB, the results of linkage analysis (Table 5.4) and β1,4-galactosyltransferase incubation (Figure 5.8) have confirmed that some glycans, such as m/z 2489, 2850 and 3212, possess bisecting GlcNAc. N-glycans with the same m/z value were observed in evCTB (Figure 5.9). There is no obvious evidence for the existence of truncated GlcNAc, for instance, the loss of two separate HexNAcs or two HexNAcs lost as a whole in the MS/MS analysis of the glycan in evCTB. In addition, it has been reported that evCTB expresses integrins (Tarrade et al., 2001; Orozco and Lewis, 2010), and integrins possess bisected N-glycans (Isaji et al., 2010; Zhang et al., 2015). All these imply that bisecting GlcNAc is also present in evCTB. Therefore I can reasonably conclude that bisecting GlcNAc is present in CTB, STB and evCTB. The bisecting GlcNAc-containing N-glycans accounted for approximately 39% (mean value of four samples) of the complex glycans in CTB and approximately 32% (mean value of four samples) of the complex glycans in STB. The bisecting GlcNAc-containing N-glycans accounted for approximately 19% (mean value of seven samples) of the complex glycans in evCTB. This shows that there are relative more abundant bisecting GlcNAc-containing N-glycans in CTB and STB.

Fucosylation was also compared among these trophoblasts, and no variation was observed (Table 5.3 and Table 5.5).

5.4.1. The potential function of NeuAc

The α2,3-linked NeuAc is formed between the C2 position of NeuAc and C3 position of Gal under the catalysis of α2,3-sialyltransferase. As shown in all the MS spectra, NeuAc is typically found to be the terminal residue. Its exposed terminal position in the carbohydrate chains make it function as a protective shield for the penultimate galactose residue. In

addition to protecting the monosaccharide, sialic acid can protect cells. For instance, erythrocytes are masked by a layer of sialic acid molecules which can prevent erythrocytes from being degraded by macrophages via phagocytosis. Sialic acids will be gradually removed from the surface by serum sialidase and spontaneous chemical hydrolysis during the cell life span which is usually 120-day. Finally, erythrocytes without sialic acid layer will be degraded as the surface uncapped galactose residues present signals for the degradation (Bratosin et al., 1995). In CTB, STB and evCTB, more than 50% of the complex N-glycans are sialylated (see section 5.3.1.1 and 5.3.2.1), it is therefore possible that NeuAcs in CTB, STB and evCTB may function in the same way to protect the trophoblasts against macrophages, especially when evCTB invades the maternal decidua and myometrium as macrophages represent approximately 20% of the decidual leukocytes at the maternal-foetal interface (Lessin et al., 1988; Houser et al., 2011).

It is reported that STB and CTB do not express HLA class I and II antigens (Murphy et al., 2004; Clark and Schust, 2013; Hunt et al., 1987; Hunt et al., 2005; Apps et al., 2009). The absence of HLA antigens would make these trophoblasts potential targets for NK cells. However, CTB and STB are resistant to NK cells due to the presence bisecting GlcNAc on their cell surfaces. Moreover, researchers found that NeuAc also contributes to this resistance: the siglec-7 (sialic acid-binding Ig-like lectin-7) is expressed predominantly on NK cells (Nicoll et al., 1999). It functions as an inhibitory receptor on human NK cells (Crocker and Varki, 2001), therefore interactions between siglec-7 and its ligands, sialic acid, can influence NK cell activity. This implies that the α2,3-linked NeuAc in CTB and STB is the potential ligand for this siglec. Siglec-7 shows similar binding affinity towards terminal α2,3-linked and α2,6-linked NeuAc (Yamaji et al., 2002).

In addition, due to the negative charge of NeuAc, it is involved in the repulsion phenomena between cells. For instance, the negative charge repulsion is a possible explanation for tumor cell detachment from the tumour mass and entering the blood stream (Fuster and Esko, 2005). This also provides possible explanation for evCTB movement from foetal part into the maternal decidua.

5.4.2.   The potential biological roles of Lewis structures

Lewis, the term, derives from a family, the members of which suffered from an erythrocyte incompatibility that helped in the discovery of blood group substances. The blood group antigens are composed of a related series of glycans possessing α1,3-linked or α1,4-linked Fuc residues.

Although it had been reported that Lewis B structure is present on human eggs (Lucas et al., 1994), it could not be confirmed physicochemically as the structure was not observed in the MALDI-MS spectra of human zona pellucida (Pang et al., 2011). However, it has been physicochemically confirmed that Lewis X/Lewis Y terminated N-glycans are expressed on human sperm (Pang et al., 2007).

The Lewis X/Lewis Y terminated N-glycans that are expressed on human sperm are thought to suppress immune responses directed against the gamete (Pang et al., 2007; Clark, 2014). This proposal was made based on a study of phase variation of *Helicobacter pylori* which is a human gastric pathogen. Variants of *H. pylori* that express Lewis X and Lewis Y structures on their lipopolysaccharides can regulate antigen related responses. This regulation is achieved via the interaction between these lipopolysaccharides and DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin) on dendritic cells; this regulation has an influence on T helper cell 1/ T helper cell 2 balance (Bergman et al., 2004). However, variants of *H. pylori* that do not express Lewis structures cause strong T helper cell 1 response, as a result of which peptic ulcer may occur (Bergman et al., 2004; Bergman et al., 2006). These suggest that the presence of Lewis structure on lipopolysaccharides play an essential role in triggering the immune tolerance. This triggering of immune tolerance may not be limited in *H. pylori*. Indeed, schistosomes, a type of helminthic parasites, express both Lewis X and a pseudo-Lewis Y that can also interact with DC-SIGN (Meyer et al., 2005). Thus, it is possible that human sperm could follow the same pathway that *H. pylori* employs to trigger immune tolerance as the sperm also expresses Lewis structures. Similarly, this pathway may be passed on to CTB, STB and evCTB with the expression of Lewis structures.

The oligosaccharide sialyl-Lewis X can function as a ligand for L-, E- and P-selectin which are important cell adhesion molecules (Phillips et al., 1990; Polley et al., 1991; Foxall et al., 1992; Julien et al., 2011). The interactions between the adhesive molecules from foetus and mother are essential for the successful implantation. It is reported that sialyl-Lewis X/L-

selectin adhesion system mediates adhesion in implantation model (Liu et al., 2011). This implies that the importance of sialyl-Lewis X on CTB and STB. It is known that leukocyte tethering and rolling is related to the interaction between the selectins and sialyl-Lewis X: selectins promote the tethering and subsequent rolling of leukocytes along the vessel wall via the sialyl-Lewis X, and this is a prelude to the extravasation into the underlying tissue (Somers et al., 2000). Since evCTB expresses sialyl-Lewis X, it is possible that sialyl-Lewis X could play an important role when evCTB are extravasated and spread out in the maternal part.

5.4.3.   The potential function of polyLacNAc

The signals of the polyLacNAc containing glycans become more obvious after sialidase S digestion, which implies that the real amount of these glycans is substantial.

It is probably easier for polyLacNAc to be involved in galectin binding as single LacNAc may be limited by its length.   Galectins are a subfamily of glycan-binding proteins that possess carbohydrate recognition domain (CRD) with specific affinity for β-galactoside (Barondes et al., 1994). Galectins are able to act both extracellularly and intracellularly to control cell fate: extracellularly, they cross link glycan ligands to transmit signals that result in death or affect other signals influencing cell fate; intracellularly, they direct regulate signals controlling cell fate (Hernandez and Baum, 2002; Yang et al., 2008).

Galectins was not identified in the eutherian uterus until 1999. From then on galectins started to attract a lot of attention due to the fact that human placental protein (PP13) is a galectin (Than et al., 1999). It has been confirmed that PP13 is expressed on STB (Than et al., 2004). Therefore, CTB probably also expresses PP13 as STB is originally differentiated from CTB. PP13 can induce apoptosis of T cells *in vitro* and kill macrophages in the maternal decidua *in situ* (Than et al., 2014). However, the underlying mechanism is still not clear. It is possible that CTB and STB could kill the immune cells via the expressed PP13 and thus suppress immune responses. In addition, it is reported that in humans 16 different galectin genes are expressed at the maternal-foetal interface (Than et al., 2009). These suggest important immune functions of galectins. Indeed, in mice galectin-3 has been reported to form lattices on cell surfaces by cross-linking glycoproteins on T cells that possess polyLacNAc moiety in

their tri- and tetraantennary N-glycans and thus blocking immune responses (Demetriou et al., 2001). Therefore, in humans the blocking is probably achieved via erecting a "galectin shield" around CTB and STB.

In addition, overexpression of polyLacNAc is one of the most common types of aberrant glycosylation observed in human cancers (Hakomori, 1996). The reason why cancer cells can evade human immune system might be related to the polyLacNAc overexpression. Compared to CTB and STB, polyLacNAc in evCTB is overexpressed (Figure 5.5 and Figure 5.11), and this could enable evCTB to evade the maternal immune system.

5.4.4. The potential function of bisecting GlcNAc

Bisecting GlcNAc is a structure that is also expressed on sperm (see section 1.8.3). Sperm expresses bisected N-glycans to inhibit NK cell cytotoxicity (see section 1.8.3). As explained in section 5.4.1 CTB and STB are the targets for NK cells due to the absence of HLA antigens on their surface, and CTB and STB also express bisected N-glycans (see sections 5.3.1.7 and 5.3.1.8). It is possible that these bisected N-glycans are expressed for the same purpose to inhibit NK cell cytotoxicity. However, the mechanism underlying the inhibition is not clear.

All these results provide more support for the human fetoembryonic defense system hypothesis which links the expression of functional glycan groups to the protection of gametes and foetus *in utero*.

Chapter 6

Concluding remarks

6. Concluding remarks

This thesis documents three separate but similar research projects. The first project (Chapter 3) used a glycomic strategy to show differences in glycan profiles between HeLa cells and NHDF. The second project (Chapter 4) employed the same strategy to identify difference in glycosylation among two healthy controls, three CMS patients caused by *GFPT1* mutations, one CMS patient caused by *DOK 7* mutation, one myopathy patient caused by *MTND5* mutation, one LGMD2A patient and one Pompe disease patient, though the quality of the results are good no significant link was found. The third project (Chapter 5) employed highly sensitive mass spectrometry and various enzymatic digestions to demonstrate the presence of α2,3-linked NeuAc, Lewis structures, sialyl-Lewis X, polyLacNAc and bisecting GlcNAc in CTB, STB and evCTB. These three projects are all about investigating glycosylation using MS, especially the first two projects which are focusing on glycosylation investigation between normal cells and abnormal cells.

Several contributions to the field of glycobiology have been made during the research:

1. It reports, for the first time, the glycolipid glycan profile of *in vitro* cultured normal human dermal fibroblasts (NHDF). Dermal fibroblasts are usually used in wound healing studies (Sorrell and Caplan, 2004; Froget et al., 2003). In addition, they are also important models for tissue regeneration and engineering studies and cancer research (Wong et al., 2007; Kalluri and Zeisberg, 2006). Glycolipid glycosylation has attracted a lot of attention due to its importance in wound healing and developing cancer therapeutic approaches (Wang et al., 2014; Daniotti et al., 2013). Therefore, knowing the glycolipid glycan profile of the fibroblasts might shed light on these research areas.

2. It also for the first time reports the N-glycan profile of *in vitro* cultured human myoblasts and myotubes differentiating from *in vitro* cultured myoblasts, the O-glycan profile of *in vitro* cultured human myotubes. Myoblasts and myotubes are precursors of mature muscle cells, so glycan profile of myoblasts and myotubes would be similar to that of muscle cells. Muscle cells are essential in musculoskeletal system (locomotor system) as they receive signals from motorneurons and then execute contraction (Bentzinger et al., 2012; Abmayr and Pavlath, 2012; Senderek et al., 2011). To achieve receiving signal, muscle cell surface receptors, such as acetylcholine receptors, are required. Acetylcholine receptors are glycoproteins and are

concentrated at neuromuscular synapses. In addition, many vital neuromuscular proteins, such as agrin and dystroglycan, are also glycosylated (Sanes and Cheney, 1982; Conti-Tronconi and Raftery, 1982; Senderek et al., 2011; Martin, 2002). However, the structures of the glycans on these glycoproteins are not clear. Therefore, knowing the glycan profile of myoblasts and myotubes could provide possible structures for those glycans, and thus explain how these glycans function.

3. It is the first report of N-glycan profile of *in vitro* cultured human CTB, STB and more authentic evCTB. N-glycan structures that expressed on human gametes could be found in the N-glycan profile of the three types of trophoblasts, which supports the human fetoembryonic defense system hypothesis. The preeclampsia that arises in the placenta is hypothesized to have an immune etiology (Laresgoiti-Servitje et al., 2010). Comparing the N-glycan profile of trophoblasts from normal pregnant woman with preeclampsia patient may provide an explanation for the etiology.

Chapter 3 describes the comparative glycomic analysis of HeLa cells and NHDF cells. The aim of the project was to find whether the cytotoxicity difference was caused by the difference in the glycan patterns of the cell surface. Not only were the glycans from glycoproteins of HeLa cells and NHDF analysed, but also the glycolipid glycans. The glycomic results showed a clear correlation with the results from Van Damme and Shang, which indicates the importance of protein-carbohydrate interactions for the proteins entering the cells. It might provide valuable insights into delivering drugs into cancer cells. For instance, the glycomic results show that HeLa cells express sialyl Lewis X structure while NHDF cells do not, therefore the sialyl Lewis X structure could be a potential target for the anticancer drug and thus enhance the deliver efficiency of the drug and avoid normal cells being killed by the drug. Simultaneously, this project provides the reader with a general understanding of currently commonly used glycomic methodology handling glycans derived from both glycoproteins and glycolipids and how mass spectrometry data are analysed and annotated. However, there are still some unsolved problems which are associated with this research project; for instance, the mechanism of the cytotoxicity of the proteins under study still needs to be uncovered. In addition, if there were sulphated glycans in HeLa and NHDF cells, their information was missing due to the permethylation strategy employed. The permethylation of sulphated glycans needs to be carried at 4 °C due to the thermal instability

of the bond(s) formed between the sulphate group(s) and the glycan, and the reaction time will be approximately 3 hours (Khoo and Yu, 2010). Following the permethylation strategy employed (see section 2.2.8), sulphated glycans will probably not be permethylated, and thus their information was lost.

Chapter 4 describes the comparative glycomic analyses of *in vitro* cultured myoblasts and myotubes from healthy controls, CMS patients caused by *GFPT1* mutations and other muscular disease patients. GFPT1 is the first enzyme of the hexosamine biosynthesis pathway (see section 1.7.1) and it catalyses the reaction providing the precursor for UDP-GlcNAc synthesis. In humans, it has a redundancy, GFPT2 which is proposed to have the same function as *GFPT1* due to the fact that their amino acid sequences share approximately 75% identity (Oki et al., 1999; Nakaishi et al., 2009). However, it is reported that GFPT2 is mainly expressed in the central nervous system (Oki et al., 1999). Therefore, it is not necessary to consider GFPT2 here. Mutations in *GFPT1* were expected to cause N-glycan branching variations and have an influence on protein glycosylation. However, the results showed that there was little difference among the glycan profile of healthy controls, *GFPT1* mutations caused CMS patients and other patients. This does not rule out the possibility that individual glycoproteins have altered glycosylation. The only way to look for this possibility is to target candidates and purify them. Current glycoprotein isolation and purification techniques cannot provide enough material for the glycomic analysis of glycoproteins of interest, which means that glycoprotein sample quantity limitations remain a challenge for glycomic studies. Therefore, these techniques still need to be further improved or developed. Future experiments aimed at the investigation of O-GlcNAcylation in the cells may enable the discovery of alternations in the glycosylation.

Chapter 5 identifies the presence of several functionally important glycan structures (Lewis X, Lewis Y, sialyl-Lewis X, polyLacNAc and bisecting GlcNAc) in CTB, STB and evCTB. This chapter is a test of the human fetoembryonic defense system hypothesis (hu-FEDS). The results support the concept that the glycosylation patterns in these three types of trophoblasts could play a role in the suppression of immune responses at the pregnant uterus, and thus supports the hypothesis. The limitations of sample quantity remain a challenge in this project as usually multiple experiments are required on the same sample to provide full structural

information. In addition, more research is required to determine the function of these glycan structures.

The NeuAc on trophoblasts (CTB, STB and evCTB) is proposed to contribute to NK cell resistance (see section 5.4.1), this could be tested by removing all NeuAc on CTB, STB and evCTB using sialidase A and then checking their NK cell resistance.

A mouse model could be set up to test the potential function of Lewis structures mentioned in section 5.4.2. Since Lewis structures have a common point: they all have an extra fucose which is either α1,3-linked or α1,4-linked. These linkages are different from the core fucose linkage which is α1,6-linked. Therefore, it is possible to knock down the expression of α1,3 and α1,4-fucosyltransferases (Gallego et al., 2003; Prohaska et al., 1978; Witte et al., 1997) and then compare the embryo development in the knock down mouse and normal mouse. It is expected that the knock down mouse would suffer abortion which could be caused by immune responses.

PolyLacNAc is proposed to be easier than LacNAc to be involved in galectin binding (see section 5.4.3), this can be tested via a control experiment: two batches of CTB (or STB) cells need to be prepared, one batch will be treated with endo-β-galactosidase which is an enzyme hydrolysing internal β1,4 galactosidic linkage in this favoured repeating unit [GlcNAcβ1,3Galβ1,4]$_n$ (Scudder et al., 1983) (see section 2.2.7.4 for details), then the cells were extracted, washed and treated with UDP-Gal and β1,4-galactosyltransferase (see section 2.2.7.3 for details); the other will not be treated. After that, same amount of fluorescence labelled galectin (e.g. galectin-3) will be added to the cells. The cells and labelled galectin will be incubated. After incubation iced cold PBS will be used to wash away the isolated galectin. The cells will be checked using fluorescence microscope. It is expected to see that the non-treated cells have stronger fluorescent signals than the cells treated with endo-β-galactosidase and β1,4-galactosyltransferase as the N-glycans in the untreated cells have polyLacNAc while the N-glycans in the treated cells only have LacNAc.

An experiment can be set up to test the potential function of bisecting GlcNAc in the trophoblasts mentioned in section 5.4.4.. Before the experiment, it is necessary to know that erythroagglutinating phytohemagglutinin (E-PHA) lectin preferentially binds to bisecting G1cNAc in N-glycans (Cummings and Kornfeld, 1982; Yoshimura et al., 1996). Firstly, CTB cells will be cultured with E-PHA (group 1) or cultured without E-PHA (group 2). In group 1, E-PHA will bind to CTB cells due to the fact that the cells express bisecting GlcNAc

containing N-glycans. Then cells from group 1 and group 2 will be resuspended and mixed with NK cells. Since the bisecting GlcNAc in group 1 is cryptic due to the presence of E-PHA, CTB cells should be killed by NK cells. While in group 2 the bisecting GlcNAc will probably prevent CTB cells being killed.

Although the three projects documented are separate, they all demonstrate the biomedical importance or potential biomedical importance of glycosylation.

It is hoped that information obtained from the first project will be helpful for seeking cervical cancer treatment. Over the last thirty years, RIPs have been applied in the construction of targeted drugs against cancer cells. These drugs are usually designed to possess two portions, one portion (lectin portion) is used for binding to cancer cells and the other (toxin portion) is used for killing them (Tejero et al., 2015; Kreitman, 2006). If the drug can bind to HeLa cells specifically, then only HeLa cells will be killed. To achieve the specific binding, it is necessary to know surface glycan structures of HeLa cells and then design the corresponding lectin portion for the drug. The glycan profiles of HeLa cells in Chapter 3 can provide clues for the lectin design. Similarly, when design lectin portion for targeted drugs against other cancers, it would be better to perform glycomic analyses of the corresponding cancer cells as the glycan profiles would be useful for the lectin design. Information obtained from the second project has ruled out the possibility that *GFPT1* mutations would cause global glycan alterations in *in vitro* cultured myoblasts and myotubes. This provides further impetus to elucidate the underlying mechanism of how *GFPT1* mutations cause CMS as future researchers will only need to focus on some particular glycoproteins, such as acetylcholine receptor and agrin. In this situation, other techniques, such lectin binding and immunoblot, could be used to investigate the glycosylation variation of those proteins of interest. For instance, RL2 antibody which binds to single GlcNAc at serine or threonine residues can be applied to check the variation of O-GlcNAcylation in glycoproteins (Tashima and Stanley, 2014).

Results obtained from the third project provide support for the human fetoembryonic defense system hypothesis which links the expression of functional glycan groups to the protection of gametes and foetus *in utero*. It is possible that when implantation occurs, the same protection system that was used to protect sperm and eggs are incorporated into trophoblasts. Therefore, this may provide a possible explanation for some of those couples who cannot make babies: their gametes may have deficient glycosylation, and thus the trophoblasts also have deficient

glycosylation, which could not induce immune tolerance. This can be checked via comparing the N-glycan profiles of their gametes to the normal couples.

References

ABMAYR, S. M. & PAVLATH, G. K. 2012. Myoblast fusion: lessons from flies and mice. *Development,* 139**,** 641-656.

AEBI, M., BERNASCONI, R., CLERC, S. & MOLINARI, M. 2010. N-glycan structures: recognition and processing in the ER. *Trends Biochem Sci,* 35**,** 74-82.

ALSTON, C. L., MORAK, M., REID, C., HARGREAVES, I. P., POPE, S. A., LAND, J. M., HEALES, S. J., HORVATH, R., MUNDY, H. & TAYLOR, R. W. 2010. A novel mitochondrial MTND5 frameshift mutation causing isolated complex I deficiency, renal failure and myopathy. *Neuromuscul Disord,* 20**,** 131-5.

ANTONOPOULOS, A., DEMOTTE, N., STROOBANT, V., HASLAM, S. M., VAN DER BRUGGEN, P. & DELL, A. 2012. Loss of effector function of human cytolytic T lymphocytes is accompanied by major alterations in N- and O-glycosylation. *J Biol Chem,* 287**,** 11240-51.

APPS, R., MURPHY, S. P., FERNANDO, R., GARDNER, L., AHAD, T. & MOFFETT, A. 2009. Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined using single antigen beads to characterize allotype specificities of anti-HLA antibodies. *Immunology,* 127**,** 26-39.

APWEILER, R., HERMJAKOB, H. & SHARON, N. 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica Et Biophysica Acta-General Subjects,* 1473**,** 4-8.

ARNOLD, J. N., WORMALD, M. R., SIM, R. B., RUDD, P. M. & DWEK, R. A. 2007. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annual Review of Immunology,* 25**,** 21-50.

BABU, P., NORTH, S. J., JANG-LEE, J., CHALABI, S., MACKERNESS, K., STOWELL, S. R., CUMMINGS, R. D., RANKIN, S., DELL, A. & HASLAM, S. M. 2009. Structural characterisation of neutrophil glycans by ultra sensitive mass spectrometric glycomics methodology. *Glycoconj J,* 26**,** 975-86.

BANDA, K., GREGG, C. J., CHOW, R., VARKI, N. M. & VARKI, A. 2012. Metabolism of vertebrate amino sugars with N-glycolyl groups: mechanisms underlying gastrointestinal incorporation of the non-human sialic acid xeno-autoantigen N-glycolylneuraminic acid. *J Biol Chem,* 287**,** 28852-64.

BANNWARTH, S., PROCACCIO, V., LEBRE, A. S., JARDEL, C., CHAUSSENOT, A., HOARAU, C., MAOULIDA, H., CHARRIER, N., GAI, X., XIE, H. M., FERRE, M., FRAGAKI, K., HARDY, G., MOUSSON DE CAMARET, B., MARLIN, S., DHAENENS, C. M., SLAMA, A., ROCHER, C., PAUL BONNEFONT, J., ROTIG, A., AOUTIL, N., GILLERON, M., DESQUIRET-DUMAS, V., REYNIER, P., CERESUELA, J., JONARD, L., DEVOS, A., ESPIL-TARIS, C., MARTINEZ, D., GAIGNARD, P., LE QUAN SANG, K. H., AMATI-BONNEAU, P., FALK, M. J., FLORENTZ, C., CHABROL, B., DURAND-ZALESKI, I. & PAQUIS-FLUCKLINGER, V. 2013. Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders. *J Med Genet,* 50**,** 704-14.

BARBER, M., BORDOLI, R. S., GARNER, G. V., GORDON, D. B., SEDGWICK, R. D., TETLER, L. W. & TYLER, A. N. 1981a. Fast-atom-bombardment mass spectra of enkephalins. *Biochem J,* 197**,** 401-4.

BARBER, M., BORDOLI, R. S., SEDGWICK, R. D. & TYLER, A. N. 1981b. Fast Atom Bombardment of Solids (Fab) - a New Ion-Source for Mass-Spectrometry. *Journal of the Chemical Society-Chemical Communications***,** 325-327.

BARBER, M., BORDOLI, R. S., SEDGWICK, R. D. & TYLER, A. N. 1981c. Fast Atom Bombardment of Solids as an Ion-Source in Mass-Spectrometry. *Nature,* 293**,** 270-275.

BARBER, M., BORDOLI, R. S., SEDGWICK, R. D., TYLER, A. N. & BYCROFT, B. W. 1981d. Fast atom bombardment mass spectrometry of bleomycin A2 and B2 and their metal complexes. *Biochem Biophys Res Commun,* 101**,** 632-8.

BARBIERI, L., CIANI, M., GIRBES, T., LIU, W. Y., VAN DAMME, E. J., PEUMANS, W. J. & STIRPE, F. 2004. Enzymatic activity of toxic and non-toxic type 2 ribosome-inactivating proteins. *FEBS Lett,* 563**,** 219-22.

BARNES, J. H. & HIEFTJE, G. M. 2004. Recent advances in detector-array technology for mass spectrometry. *International Journal of Mass Spectrometry,* 238**,** 33-46.

BARONDES, S. H., CASTRONOVO, V., COOPER, D. N., CUMMINGS, R. D., DRICKAMER, K., FEIZI, T., GITT, M. A., HIRABAYASHI, J., HUGHES, C., KASAI, K. & ET AL. 1994. Galectins: a family of animal beta-galactoside-binding lectins. *Cell,* 76**,** 597-8.

BARRY, C. S., COCINERO, E. J., CARCABAL, P., GAMBLIN, D. P., STANCA-KAPOSTA, E. C., REMMERT, S. M., FERNANDEZ-ALONSO, M. C., RUDIC, S., SIMONS, J. P. & DAVIS, B. G. 2013. 'Naked' and hydrated conformers of the conserved core pentasaccharide of N-linked glycoproteins and its building blocks. *J Am Chem Soc,* 135**,** 16895-903.

BASU, S., KAUFMAN, B. & ROSEMAN, S. 1968. Enzymatic synthesis of ceramide-glucose and ceramide-lactose by glycosyltransferases from embryonic chicken brain. *J Biol Chem,* 243**,** 5802-4.

BATEMAN, A. C., KARAMANSKA, R., BUSCH, M. G., DELL, A., OLSEN, C. W. & HASLAM, S. M. 2010. Glycan analysis and influenza A virus infection of primary swine respiratory epithelial cells: the importance of NeuAc{alpha}2-6 glycans. *J Biol Chem,* 285**,** 34016-26.

BATTELLI, M. G., CITORES, L., BUONAMICI, L., FERRERAS, J. M., DE BENITO, F. M., STIRPE, F. & GIRBES, T. 1997. Toxicity and cytotoxicity of nigrin b, a two-chain ribosome-inactivating protein from Sambucus nigra: comparison with ricin. *Arch Toxicol,* 71**,** 360-4.

BAUSE, E. & LEGLER, G. 1981. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis. *Biochem J,* 195**,** 639-44.

BAX, M., GARCIA-VALLEJO, J. J., JANG-LEE, J., NORTH, S. J., GILMARTIN, T. J., HERNANDEZ, G., CROCKER, P. R., LEFFLER, H., HEAD, S. R., HASLAM, S. M., DELL, A. & VAN KOOYK, Y. 2007. Dendritic cell maturation results in pronounced changes in glycan expression affecting recognition by siglecs and galectins. *J Immunol,* 179**,** 8216-24.

BEAVIS, R. C. & CHAIT, B. T. 1991. Velocity Distributions of Intact High Mass Polypeptide Molecule Ions Produced by Matrix Assisted Laser Desorption. *Chemical Physics Letters,* 181**,** 479-484.

BEESON, D., HIGUCHI, O., PALACE, J., COSSINS, J., SPEARMAN, H., MAXWELL, S., NEWSOM-DAVIS, J., BURKE, G., FAWCETT, P., MOTOMURA, M., MULLER, J. S.,

LOCHMULLER, H., SLATER, C., VINCENT, A. & YAMANASHI, Y. 2006. Dok-7 mutations underlie a neuromuscular junction synaptopathy. *Science,* 313**,** 1975-8.

BELAYA, K., FINLAYSON, S., SLATER, C. R., COSSINS, J., LIU, W. W., MAXWELL, S., MCGOWAN, S. J., MASLAU, S., TWIGG, S. R., WALLS, T. J., PASCUAL PASCUAL, S. I., PALACE, J. & BEESON, D. 2012. Mutations in DPAGT1 cause a limb-girdle congenital myasthenic syndrome with tubular aggregates. *American Journal of Human Genetics,* 91**,** 193-201.

BENIRSCHKE, K. 1994. Anatomical Relationship between Fetus and Mother. *Fetal Cells in Maternal Blood: Prospects for Noninvasive Prenatal Diagnosis,* 731**,** 9-20.

BENIRSCHKE, K. 1998. Remarkable placenta. *Clin Anat,* 11**,** 194-205.

BENNETT, E. P., MANDEL, U., CLAUSEN, H., GERKEN, T. A., FRITZ, T. A. & TABAK, L. A. 2012. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology,* 22**,** 736-56.

BENTZINGER, C. F., WANG, Y. X. & RUDNICKI, M. A. 2012. Building Muscle: Molecular Regulation of Myogenesis. *Cold Spring Harbor Perspectives in Biology,* 4.

BERG, J. M., TYMOCZKO, J. L. & STRYER, L. 2002a. Section 8.4 The Michaelis-Menten Model Accounts for the Kinetic Properties of Many Enzymes. *Biochemistry.* 5th edition ed. New York: W H Freeman.

BERG, J. M., TYMOCZKO, J. L. & STRYER, L. 2002b. Section 11.4: Lectins Are Specific Carbohydrate-Binding Proteins. *Biochemistry.* 5th edition ed. New York: W H Freeman.

BERGAMIN, E., HALLOCK, P. T., BURDEN, S. J. & HUBBARD, S. R. 2010. The cytoplasmic adaptor protein Dok7 activates the receptor tyrosine kinase MuSK via dimerization. *Mol Cell,* 39**,** 100-9.

BERGMAN, M., DEL PRETE, G., VAN KOOYK, Y. & APPELMELK, B. 2006. Helicobacter pylori phase variation, immune modulation and gastric autoimmunity. *Nat Rev Microbiol,* 4**,** 151-9.

BERGMAN, M. P., ENGERING, A., SMITS, H. H., VAN VLIET, S. J., VAN BODEGRAVEN, A. A., WIRTH, H. P., KAPSENBERG, M. L., VANDENBROUCKE-GRAULS, C. M., VAN KOOYK, Y. & APPELMELK, B. J. 2004. Helicobacter pylori modulates the T helper cell 1/T helper cell 2 balance through phase-variable interaction between lipopolysaccharide and DC-SIGN. *J Exp Med,* 200**,** 979-90.

BERGSTROM, K. S. B. & XIA, L. J. 2013. Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology,* 23**,** 1026-1037.

BERN, M., BRITO, A. E., PANG, P. C., REKHI, A., DELL, A. & HASLAM, S. M. 2013. Polylactosaminoglycan glycomics: enhancing the detection of high-molecular-weight N-glycans in matrix-assisted laser desorption ionization time-of-flight profiles by matched filtering. *Mol Cell Proteomics,* 12**,** 996-1004.

BILLINGHAM, R. E., BRENT, L. & MEDAWAR, P. B. 1956. The antigenic stimulus in transplantation immunity. *Nature,* 178**,** 514-9.

BLOW, N. 2009. Glycobiology: A spoonful of sugar. *Nature,* 457**,** 617-622.

BOLTON, A. E., POCKLEY, A. G., CLOUGH, K. J., MOWLES, E. A., STOKER, R. J., WESTWOOD, O. M. R. & CHAPMAN, M. G. 1987. Identification of Placental Protein-14 as an Immunosuppressive Factor in Human-Reproduction. *Lancet,* 1**,** 593-595.

BRATOSIN, D., MAZURIER, J., DEBRAY, H., LECOCQ, M., BOILLY, B., ALONSO, C., MOISEI, M., MOTAS, C. & MONTREUIL, J. 1995. Flow cytofluorimetric analysis of young and senescent human erythrocytes probed with lectins. Evidence that sialic acids control their life span. *Glycoconj J,* 12**,** 258-67.

BROCKHAUSEN, I. 1999. Pathways of O-glycan biosynthesis in cancer cells. *Biochim Biophys Acta,* 1473**,** 67-95.

BROCKHAUSEN, I., YANG, J. M., BURCHELL, J., WHITEHOUSE, C. & TAYLOR-PAPADIMITRIOU, J. 1995. Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. *Eur J Biochem,* 233**,** 607-17.

BUSCH, K. L. 1995. Desorption Ionization Mass-Spectrometry. *Journal of Mass Spectrometry,* 30**,** 233-240.

CALVETE, J. J. 2014. The expanding universe of mass analyzer configurations for biological analysis. *Methods Mol Biol,* 1072**,** 61-81.

CAMERON, A. E. & EGGERS, D. F. 1948. An Ion "Velocitron". *Rev. Sci. Instrum.,* 19.

CERONI, A., MAASS, K., GEYER, H., GEYER, R., DELL, A. & HASLAM, S. M. 2008. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of Glycans. *Journal of Proteome Research,* 7**,** 1650-1659.

CHAI, W. G., HOUNSELL, E. F., CASHMORE, G. C., ROSANKIEWICZ, J. R., BAUER, C. J., FEENEY, J., FEIZI, T. & LAWSON, A. M. 1992. Neutral oligosaccharides of bovine submaxillary mucin. A combined mass spectrometry and 1H-NMR study. *Eur J Biochem,* 203**,** 257-68.

CHU, V. C. & WHITTAKER, G. R. 2004. Influenza virus entry and infection require host cell N-linked glycoprotein. *Proc Natl Acad Sci U S A,* 101**,** 18153-8.

CITORES, L., FERRERAS, J. M., MUNOZ, R., BENITEZ, J., JIMENEZ, P. & GIRBES, T. 2002. Targeting cancer cells with transferrin conjugates containing the non-toxic type 2 ribosome-inactivating proteins nigrin b or ebulin l. *Cancer Lett,* 184**,** 29-35.

CLARK, G. F. 2014. The role of glycans in immune evasion: the human fetoembryonic defence system hypothesis revisited. *Mol Hum Reprod,* 20**,** 185-99.

CLARK, G. F., DELL, A., MORRIS, H. R., PATANKAR, M. S. & EASTON, R. L. 2001. The species recognition system: a new corollary for the human fetoembryonic defense system hypothesis. *Cells Tissues Organs,* 168**,** 113-21.

CLARK, G. F., OEHNINGER, S., PATANKAR, M. S., KOISTINEN, R., DELL, A., MORRIS, H. R., KOISTINEN, H. & SEPPALA, M. 1996. A role for glycoconjugates in human development: the human feto-embryonic defence system hypothesis. *Hum Reprod,* 11**,** 467-73.

CLARK, G. F. & SCHUST, D. J. 2013. Manifestations of immune tolerance in the human female reproductive tract. *Front Immunol,* 4**,** 26.

COHEN, M., HURTADO-ZIOLA, N. & VARKI, A. 2009. ABO blood group glycans modulate sialic acid recognition on erythrocytes. *Blood,* 114**,** 3668-76.

CONTI-TRONCONI, B. M. & RAFTERY, M. A. 1982. The nicotinic cholinergic receptor: correlation of molecular structure with functional properties. *Annual Review of Biochemistry,* 51**,** 491-530.

CORFIELD, A. P., HIGA, H., PAULSON, J. C. & SCHAUER, R. 1983. The specificity of viral and bacterial sialidases for alpha(2-3)- and alpha(2-6)-linked sialic acids in glycoproteins. *Biochim Biophys Acta,* 744**,** 121-6.

COSSINS, J., BELAYA, K., HICKS, D., SALIH, M. A., FINLAYSON, S., CARBONI, N., LIU, W. W., MAXWELL, S., ZOLTOWSKA, K., FARSANI, G. T., LAVAL, S., SEIDHAMED, M. Z., DONNELLY, P., BENTLEY, D., MCGOWAN, S. J., MULLER, J., PALACE, J., LOCHMULLER, H., BEESON, D. & CONSORTIUM, W. 2013. Congenital myasthenic syndromes due to mutations in ALG2 and ALG14. *Brain,* 136**,** 944-956.

COUTINHO, P. M., DELEURY, E., DAVIES, G. J. & HENRISSAT, B. 2003. An evolving hierarchical family classification for glycosyltransferases. *Journal of Molecular Biology,* 328**,** 307-17.

CROCKER, P. R. & VARKI, A. 2001. Siglecs in the immune system. *Immunology,* 103**,** 137-45.

CULYBA, E. K. 2012. Protein native-state stabilization by placing aromatic side chains in N-glycosylated reverse turns (February, pg 571, 2011). *Science,* 335**,** 1042-1042.

CUMMINGS, R. D. & KORNFELD, S. 1982. Characterization of the structural determinants required for the high affinity interaction of asparagine-linked oligosaccharides with immobilized Phaseolus vulgaris leukoagglutinating and erythroagglutinating lectins. *J Biol Chem,* 257**,** 11230-4.

CUMMINGS, R. D. & PIERCE, J. M. 2014. The challenge and promise of glycomics. *Chem Biol,* 21**,** 1-15.

D'ANGELO, G., CAPASSO, S., STICCO, L. & RUSSO, D. 2013. Glycosphingolipids: synthesis and functions. *FEBS J,* 280**,** 6338-53.

D'ANGELO, G., POLISHCHUK, E., DI TULLIO, G., SANTORO, M., DI CAMPLI, A., GODI, A., WEST, G., BIELAWSKI, J., CHUANG, C. C., VAN DER SPOEL, A. C., PLATT, F. M., HANNUN, Y. A., POLISHCHUK, R., MATTJUS, P. & DE MATTEIS, M. A. 2007. Glycosphingolipid synthesis requires FAPP2 transfer of glucosylceramide. *Nature,* 449**,** 62-7.

D'ANGELO, G., REGA, L. R. & DE MATTEIS, M. A. 2012. Connecting vesicular transport with lipid synthesis: FAPP2. *Biochim Biophys Acta,* 1821**,** 1089-95.

DANIOTTI, J. L., VILCAES, A. A., TORRES DEMICHELIS, V., RUGGIERO, F. M. & RODRIGUEZ-WALKER, M. 2013. Glycosylation of glycolipids in cancer: basis for development of novel therapeutic approaches. *Front Oncol,* 3**,** 306.

DAS, M. K., SHARMA, R. S. & MISHRA, V. 2012. Induction of apoptosis by ribosome inactivating proteins: importance of N-glycosidase activity. *Appl Biochem Biotechnol,* 166**,** 1552-61.

DELL, A. 1987. Fab-Mass Spectrometry of Carbohydrates. *Advances in Carbohydrate Chemistry and Biochemistry,* 45**,** 19-72.

DELL, A. 1990. Preparation and Desorption Mass-Spectrometry of Permethyl and Peracetyl Derivatives of Oligosaccharides. *Methods in Enzymology,* 193**,** 647-660.

DELL, A., GALADARI, A., SASTRE, F. & HITCHEN, P. 2010. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int J Microbiol,* 2010**,** 148178.

DELL, A. & MORRIS, H. R. 1974. New observations on the fragmentation properties of peptides under electron impact mass spectrometry. *Biochem Biophys Res Commun,* 61**,** 1125-32.

DELL, A. & MORRIS, H. R. 1982. Fast Atom Bombardment - High-Field Magnet Mass-Spectrometry of 6000 Dalton Polypeptides. *Biochemical and Biophysical Research Communications,* 106**,** 1456-1462.

DELL, A. & MORRIS, H. R. 2001. Glycoprotein structure determination mass spectrometry. *Science,* 291**,** 2351-2356.

DELL, A., MORRIS, H. R., EGGE, H., VONNICOLAI, H. & STRECKER, G. 1983a. Fast-Atom-Bombardment Mass-Spectrometry for Carbohydrate-Structure Determination. *Carbohydrate Research,* 115**,** 41-52.

DELL, A., MORRIS, H. R., LEVIN, M. D. & HECHT, S. M. 1981. Field desorption and fast atom bombardment mass spectrometry of bleomycins and their derivatives. *Biochem Biophys Res Commun,* 102**,** 730-8.

DELL, A., OATES, J. E., MORRIS, H. R. & EGGE, H. 1983b. Structure Determination of Carbohydrates and Glycosphingolipids by Fast Atom Bombardment Mass-Spectrometry. *International Journal of Mass Spectrometry and Ion Processes,* 46**,** 415-418.

DELL, A., REASON, A. J., KHOO, K. H., PANICO, M., MCDOWELL, R. A. & MORRIS, H. R. 1994. Mass-Spectrometry of Carbohydrate-Containing Biopolymers. *Guide to Techniques in Glycobiology,* 230**,** 108-132.

DEMETRIOU, M., GRANOVSKY, M., QUAGGIN, S. & DENNIS, J. W. 2001. Negative regulation of T-cell activation and autoimmunity by Mgat5 N- glycosylation. *Nature,* 409**,** 733-9.

DI GIROLAMO, F., LANTE, I., MURACA, M. & PUTIGNANI, L. 2013. The Role of Mass Spectrometry in the "Omics" Era. *Curr Org Chem,* 17**,** 2891-2905.

DOWNHAM, E., WINTERTHUN, S., NAKKESTAD, H. L., HIRTH, A., HALVORSEN, T., TAYLOR, R. W. & BINDOFF, L. A. 2008. A novel mitochondrial ND5 (MTND5) gene mutation giving isolated exercise intolerance. *Neuromuscul Disord,* 18**,** 310-4.

EGGE, H., PETER-KATALINIC, J., PAZ-PARENTE, J., STRECKER, G., MONTREUIL, J. & FOURNET, B. 1983. Carbohydrate structures of hen ovomucoid. A mass spectrometric analysis. *FEBS Lett,* 156**,** 357-62.

ELHAMMER, A. P., POORMAN, R. A., BROWN, E., MAGGIORA, L. L., HOOGERHEIDE, J. G. & KEZDY, F. J. 1993. The Specificity of Udp-Galnac - Polypeptide N-Acetylgalactosaminyltransferase as Inferred from a Database of Invivo Substrates and from the Invitro Glycosylation of Proteins and Peptides. *Journal of Biological Chemistry,* 268**,** 10029-10038.

ELLIES, L. G., TSUBOI, S., PETRYNIAK, B., LOWE, J. B., FUKUDA, M. & MARTH, J. D. 1998. Core 2 oligosaccharide biosynthesis distinguishes between selectin ligands essential for leukocyte homing and inflammation. *Immunity,* 9**,** 881-890.

ENDO, Y., MITSUI, K., MOTIZUKI, M. & TSURUGI, K. 1987. The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins. *J Biol Chem,* 262**,** 5908-12.

ENDO, Y. & TSURUGI, K. 1987. RNA N-glycosidase activity of ricin A-chain. Mechanism of action of the toxic lectin ricin on eukaryotic ribosomes. *J Biol Chem,* 262**,** 8128-30.

ENGEL, A. G., SHEN, X. M., SELCEN, D. & SINE, S. M. 2015. Congenital myasthenic syndromes: pathogenesis, diagnosis, and treatment. *Lancet Neurol,* 14**,** 420-34.

ENGELSTAEDTER, V., FLUEGEL, B., KUNZE, S., MAYR, D., FRIESE, K., JESCHKE, U. & BERGAUER, F. 2012. Expression of the carbohydrate tumour marker Sialyl Lewis A, Sialyl Lewis X, Lewis Y and Thomsen-Friedenreich antigen in normal squamous epithelium of the uterine cervix, cervical dysplasia and cervical cancer. *Histol Histopathol,* 27**,** 507-14.

FANIN, M. & ANGELINI, C. 2015. Protein and genetic diagnosis of limb girdle muscular dystrophy type 2A: The yield and the pitfalls. *Muscle Nerve,* 52**,** 163-73.

FDA 2014. Guidance for Industry: Immunogenicity Assessment for Therapeutic Protein Products. *In:* SERVICES, U. S. D. O. H. A. H. (ed.).

FERRERAS, J. M., CITORES, L., IGLESIAS, R., JIMENEZ, P. & GIRBES, T. 2011. Use of ribosome-inactivating proteins from Sambucus for the construction of immunotoxins and conjugates for cancer therapy. *Toxins (Basel),* 3**,** 420-41.

FOXALL, C., WATSON, S. R., DOWBENKO, D., FENNIE, C., LASKY, L. A., KISO, M., HASEGAWA, A., ASA, D. & BRANDLEY, B. K. 1992. The three members of the selectin receptor family recognize a common carbohydrate epitope, the sialyl Lewis(x) oligosaccharide. *J Cell Biol,* 117**,** 895-902.

FREEZE, H. H. 2006. Genetic defects in the human glycome. *Nature Reviews Genetics,* 7**,** 537-551.

FREEZE, H. H. & AEBI, M. 2005. Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr Opin Struct Biol,* 15**,** 490-8.

FREEZE, H. H., CHONG, J. X., BAMSHAD, M. J. & NG, B. G. 2014. Solving glycosylation disorders: fundamental approaches reveal complicated pathways. *Am J Hum Genet,* 94**,** 161-75.

FREEZE, H. H., EKLUND, E. A., NG, B. G. & PATTERSON, M. C. 2015. Neurological Aspects of Human Glycosylation Disorders. *Annu Rev Neurosci,* 38**,** 105-25.

FROGET, S., BARTHELEMY, E., GUILLOT, F., SOLER, C., COUDERT, M. C., BENBUNAN, M. & DOSQUET, C. 2003. Wound healing mediator production by human dermal fibroblasts grown within a collagen-GAG matrix for skin repair in humans. *Eur Cytokine Netw,* 14**,** 60-4.

FUKUDA, M., LAUFFENBURGER, M., SASAKI, H., ROGERS, M. E. & DELL, A. 1987. Structures of novel sialylated O-linked oligosaccharides isolated from human erythrocyte glycophorins. *J Biol Chem,* 262**,** 11952-7.

FUKUDA, M., SPOONCER, E., OATES, J. E., DELL, A. & KLOCK, J. C. 1984. Structure of sialylated fucosyl lactosaminoglycan isolated from human granulocytes. *J Biol Chem,* 259**,** 10925-35.

FUNAKOSHI, T., YASUDA, S., FUKASAWA, M., NISHIJIMA, M. & HANADA, K. 2000. Reconstitution of ATP- and cytosol-dependent transport of de novo synthesized ceramide to the site of sphingomyelin synthesis in semi-intact cells. *J Biol Chem,* 275**,** 29938-45.

FUSTER, M. M. & ESKO, J. D. 2005. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer,* 5**,** 526-42.

GALLEGO, R. G., DUDZIAK, G., KRAGL, U., WANDREY, C., KAMERLING, J. P. & VLIEGENTHART, J. F. G. 2003. Enzymatic synthesis of the core-2 sialyl Lewis X O-glycan on the tumor-associated MUC1a' peptide. *Biochimie,* 85**,** 275-286.

GAULT, C. R., OBEID, L. M. & HANNUN, Y. A. 2010. An overview of sphingolipid metabolism: from synthesis to breakdown. *Adv Exp Med Biol,* 688**,** 1-23.

GEBAUER, J. M., MULLER, S., HANISCH, F. G., PAULSSON, M. & WAGENER, R. 2008. O-glucosylation and O-fucosylation occur together in close proximity on the first epidermal growth factor repeat of AMACO (VWA2 protein). *J Biol Chem,* 283**,** 17846-54.

GEORGIADES, P., FERGUSON-SMITH, A. C. & BURTON, G. J. 2002. Comparative developmental anatomy of the murine and human definitive placentae. *Placenta,* 23**,** 3-19.

GILL, D. J., CLAUSEN, H. & BARD, F. 2011. Location, location, location: new insights into O-GaINAc protein glycosylation. *Trends in Cell Biology,* 21**,** 149-158.

GLISH, G. L. & VACHET, R. W. 2003. The basics of mass spectrometry in the twenty-first century. *Nat Rev Drug Discov,* 2**,** 140-50.

GOLDBERG, D., BERN, M., NORTH, S. J., HASLAM, S. M. & DELL, A. 2009. Glycan family analysis for deducing N-glycan topology from single MS. *Bioinformatics,* 25**,** 365-71.

GUERGUELTCHEVA, V., MULLER, J., DUSL, M., SENDEREK, J., OLDFORS, A., LINDBERGH, C., MAXWELL, S., COLOMER, J., MALLEBRERA, C., NASCIMENTO, A., VILCHEZ, J., MUELAS, N., KIRSCHNER, J., NAFISSI, S., KARIMINEJAD, A., NILIPOUR, Y., BOZORGMEHR, B., NAJMABADI, H., RODOLICO, C., SIEB, J., SCHLOTTER, B., SCHOSER, B., HERRMANN, R., VOIT, T., STEINLEIN, O., NAJAFI, A., URTIZBEREA, A., SOLER, D., MUNTONI, F., HANNA, M., CHAOUCH, A., STRAUB, V., BUSHBY, K., PALACE, J., BEESON, D., ABICHT, A. & LOCHMULLER, H. 2012. Congenital myasthenic syndrome with tubular aggregates caused by GFPT1 mutations. *Journal of Neurology,* 259**,** 838-850.

GUILHAUS, M. 1995. Principles and Instrumentation in Time-of-Flight Mass-Spectrometry - Physical and Instrumental Concepts. *Journal of Mass Spectrometry,* 30**,** 1519-1532.

GULERIA, I. & SAYEGH, M. H. 2007. Maternal acceptance of the fetus: true human tolerance. *J Immunol,* 178**,** 3345-51.

GUPTA, G. & SUROLIA, A. 2010. Glycosphingolipids in microdomain formation and their spatial organization. *FEBS Lett,* 584**,** 1634-41.

GUPTA, R. & BRUNAK, S. 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput***,** 310-22.

GUZMAN-ARANGUEZ, A. & ARGUESO, P. 2010. Structure and biological roles of mucin-type O-glycans at the ocular surface. *Ocul Surf,* 8**,** 8-17.

HAKOMORI, S. 1996. Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res,* 56**,** 5309-18.

HAKOMORI, S. 2003. Structure, organization, and function of glycosphingolipids in membrane. *Curr Opin Hematol,* 10**,** 16-24.

HAKOMORI, S. 2004. Glycosynapses: microdomains controlling carbohydrate-dependent cell adhesion and signaling. *An Acad Bras Cienc,* 76**,** 553-72.

HALTIWANGER, R. S., BLOMBERG, M. A. & HART, G. W. 1992. Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase. *J Biol Chem,* 267**,** 9005-13.

HANDSCHUH, K., GUIBOURDENCHE, J., TSATSARIS, V., GUESNON, M., LAURENDEAU, I., EVAIN-BRION, D. & FOURNIER, T. 2007. Human chorionic gonadotropin expression in human trophoblasts from early placenta: comparative study between villous and extravillous trophoblastic cells. *Placenta,* 28**,** 175-84.

HARRIS, R. J. & SPELLMAN, M. W. 1993. O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology,* 3**,** 219-24.

HART, G. W. 2013. Thematic minireview series on glycobiology and extracellular matrices: glycan functions pervade biology at all levels. *J Biol Chem,* 288**,** 6903.

HART, G. W., SLAWSON, C., RAMIREZ-CORREA, G. & LAGERLOF, O. 2011. Cross Talk Between O-GlcNAcylation and Phosphorylation: Roles in Signaling, Transcription, and Chronic Disease. *Annual Review of Biochemistry, Vol 80,* 80**,** 825-858.

HASHII, N., KAWASAKI, N., ITOH, S., NAKAJIMA, Y., KAWANISHI, T. & YAMAGUCHI, T. 2009. Alteration of N-glycosylation in the kidney in a mouse model of systemic lupus erythematosus: relative quantification of N-glycans using an isotope-tagging method. *Immunology,* 126**,** 336-45.

HASLAM, S. M., JULIEN, S., BURCHELL, J. M., MONK, C. R., CERONI, A., GARDEN, O. A. & DELL, A. 2008. Characterizing the glycome of the mammalian immune system. *Immunol Cell Biol,* 86**,** 564-73.

HEAD, J. R. & BILLINGHAM, R. E. 1986. Concerning the immunology of the uterus. *Am J Reprod Immunol Microbiol,* 10**,** 76-81.

HEJAZI, L., EBRAHIMI, D., HIBBERT, D. B. & GUILHAUS, M. 2009. Compatibility of electron ionization and soft ionization methods in gas chromatography/orthogonal time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry,* 23**,** 2181-2189.

HERMANS, M. M., WISSELAAR, H. A., KROOS, M. A., OOSTRA, B. A. & REUSER, A. J. 1993. Human lysosomal alpha-glucosidase: functional characterization of the glycosylation sites. *Biochem J,* 289 (Pt 3)**,** 681-6.

HERNANDEZ, F., SANCHO, J. V., IBANEZ, M., ABAD, E., PORTOLES, T. & MATTIOLI, L. 2012. Current use of high-resolution mass spectrometry in the environmental sciences. *Anal Bioanal Chem,* 403**,** 1251-64.

HERNANDEZ, J. D. & BAUM, L. G. 2002. Ah, sweet mystery of death! Galectins and control of cell fate. *Glycobiology,* 12**,** 127R-36R.

HONIG, R. E. & WOOLSTON, J. R. 1963. Laser-Induced Emission of Electrons, Ions, and Neutral Atoms from Solid Surfaces. *Applied Physics Letters,* 2**,** 138-139.

HORNING, E. C., CARROLL, D. I., DZIDIC, I., HAEGELE, K. D., LIN, S., OERTLI, C. U. & STILLWELL, R. N. 1977. Development and use of analytical systems based on mass spectrometry. *Clin Chem,* 23**,** 13-21.

HOUSER, B. L., TILBURGS, T., HILL, J., NICOTRA, M. L. & STROMINGER, J. L. 2011. Two unique human decidual macrophage populations. *J Immunol,* 186**,** 2633-42.

HUNT, J. S., ANDREWS, G. K. & WOOD, G. W. 1987. Normal trophoblasts resist induction of class I HLA. *J Immunol,* 138**,** 2481-7.

HUNT, J. S., PETROFF, M. G., MCINTIRE, R. H. & OBER, C. 2005. HLA-G and immune tolerance in pregnancy. *FASEB J,* 19**,** 681-93.

HURTADO-GUERRERO, R., DORFMUELLER, H. C. & VAN AALTEN, D. M. 2008. Molecular mechanisms of O-GlcNAcylation. *Current Opinion in Structural Biology,* 18**,** 551-557.

HUTTER, H. & DOHR, G. 1998. HLA expression on immature and mature human germ cells. *J Reprod Immunol,* 38**,** 101-22.

ICHIKAWA, S., SAKIYAMA, H., SUZUKI, G., HIDARI, K. I. & HIRABAYASHI, Y. 1996. Expression cloning of a cDNA for human ceramide glucosyltransferase that catalyzes the first glycosylation step of glycosphingolipid synthesis. *Proc Natl Acad Sci U S A,* 93**,** 4638-43.

ISAJI, T., KARIYA, Y., XU, Q., FUKUDA, T., TANIGUCHI, N. & GU, J. 2010. Functional roles of the bisecting GlcNAc in integrin-mediated cell adhesion. *Methods Enzymol,* 480**,** 445-59.

IZU, H., IZUMI, Y., KUROME, Y., SANO, M., KONDO, A., KATO, I. & ITO, M. 1997. Molecular cloning, expression, and sequence analysis of the endoglycoceramidase II gene from Rhodococcus species strain M-777. *J Biol Chem,* 272**,** 19846-50.

JAEGER, B. N. & VIVIER, E. 2012. Natural killer cell tolerance: control by self or self-control? *Cold Spring Harb Perspect Biol,* 4.

JANG-LEE, J., NORTH, S. J., SUTTON-SMITH, M., GOLDBERG, D., PANICO, M., MORRIS, H., HASLAM, S. & DELL, A. 2006. Glycomic profiling of cells and tissues by mass spectrometry: fingerprinting and sequencing methodologies. *Methods Enzymol,* 415**,** 59-86.

JI, L., BRKIC, J., LIU, M., FU, G., PENG, C. & WANG, Y. L. 2013. Placental trophoblast cell differentiation: physiological regulation and pathological relevance to preeclampsia. *Mol Aspects Med,* 34**,** 981-1023.

JIA, N., BARCLAY, W. S., ROBERTS, K., YEN, H. L., CHAN, R. W., LAM, A. K., AIR, G., PEIRIS, J. M., DELL, A., NICHOLLS, J. M. & HASLAM, S. M. 2014. Glycomic characterisation of respiratory tract tissues of ferrets: implications for its use in influenza virus infection studies. *J Biol Chem,* 289**,** 28489-28504.

JUCH, H., BLASCHITZ, A., DOHR, G. & HUTTER, H. 2012. HLA class I expression in the human placenta. *Wien Med Wochenschr,* 162**,** 196-200.

JULENIUS, K., MOLGAARD, A., GUPTA, R. & BRUNAK, S. 2005. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology,* 15**,** 153-64.

JULIEN, S., IVETIC, A., GRIGORIADIS, A., QIZE, D., BURFORD, B., SPROVIERO, D., PICCO, G., GILLETT, C., PAPP, S. L., SCHAFFER, L., TUTT, A., TAYLOR-PAPADIMITRIOU, J., PINDER, S. E. & BURCHELL, J. M. 2011. Selectin ligand sialyl-Lewis x antigen drives metastasis of hormone-dependent breast cancers. *Cancer Res,* 71**,** 7683-93.

JULIEN, S., VIDEIRA, P. A. & DELANNOY, P. 2012. Sialyl-tn in cancer: (how) did we miss the target? *Biomolecules,* 2**,** 435-66.

KAILEMIA, M. J., RUHAAK, L. R., LEBRILLA, C. B. & AMSTER, I. J. 2014. Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Anal Chem,* 86**,** 196-212.

KALLURI, R. & ZEISBERG, M. 2006. Fibroblasts in cancer. *Nat Rev Cancer,* 6**,** 392-401.

KARAS, M., BACHMANN, D. & HILLENKAMP, F. 1985. Influence of the Wavelength in High-Irradiance Ultraviolet-Laser Desorption Mass-Spectrometry of Organic-Molecules. *Analytical Chemistry,* 57**,** 2935-2939.

KARLSSON, K. A., LEFFLER, H. & SAMUELSSON, B. E. 1974. Characterization of the forssman glycolipid hapten of horse kidney by mass spectrometry. *J Biol Chem,* 249**,** 4819-23.

KARRE, K. 2002. NK cells, MHC class I molecules and the missing self. *Scand J Immunol,* 55**,** 221-8.

KAUFMANN, R., CHAURAND, P., KIRSCH, D. & SPENGLER, B. 1996. Post-source decay and delayed extraction in matrix-assisted laser desorption/ionization-reflectron time-of-flight mass spectrometry. Are there trade-offs? *Rapid Commun Mass Spectrom,* 10**,** 1199-208.

KEITH, A. N. 1999. *A History of the Mass Spectrometer* [Online]. Available: http://masspec.scripps.edu/mshistory/perspectives/knier.php [Accessed 2011-11-15 2011].

KELLY, R. W. & CRITCHLEY, H. O. D. 1997. Immunomodulation by human seminal plasma: a benefit for spermatozoon and pathogen? *Human Reproduction,* 12**,** 2200-2207.

KENT, S. 2004. Novel forms of chemical protein diversity - in nature and in the laboratory. *Current Opinion in Biotechnology,* 15**,** 607-614.

KHOO, K. H. & YU, S. Y. 2010. Mass Spectrometric Analysis of Sulfated N- and O-Glycans. *Methods in Enzymology, Vol 478: Glycomics,* 478**,** 3-26.

KIESSLING, L. L. & SPLAIN, R. A. 2010. Chemical Approaches to Glycobiology. *Annual Review of Biochemistry, Vol 79,* 79**,** 619-653.

KITAGAWA, H. & PAULSON, J. C. 1994. Differential expression of five sialyltransferase genes in human tissues. *J Biol Chem,* 269**,** 17872-8.

KLISCH, K., JEANROND, E., PANG, P. C., PICH, A., SCHULER, G., DANTZER, V., KOWALEWSKI, M. P. & DELL, A. 2008. A tetraantennary glycan with bisecting N-acetylglucosamine and the Sd(a) antigen is the predominant N-glycan on bovine pregnancy-associated glycoproteins. *Glycobiology,* 18**,** 42-52.

KNOCHENMUSS, R. 2006. Ion formation mechanisms in UV-MALDI. *Analyst,* 131**,** 966-986.

KOCHETKOV, N. K., CHIZHOV, O. S. & MOLODTSOV, N. V. 1968. Mass spectrometry of oligosaccharides. *Tetrahedron,* 24**,** 5587–5593.

KOLTER, T. 2012. Ganglioside Biochemistry. *ISRN Biochemistry,* 2012**,** 36.

KOLTER, T., PROIA, R. L. & SANDHOFF, K. 2002. Combinatorial ganglioside biosynthesis. *J Biol Chem,* 277**,** 25859-62.

KOLTER, T. & SANDHOFF, K. 1998. Recent advances in the biochemistry of sphingolipidoses. *Brain Pathology,* 8**,** 79-100.

KORNFELD, R. & KORNFELD, S. 1985. Assembly of Asparagine-Linked Oligosaccharides. *Annual Review of Biochemistry,* 54**,** 631-664.

KRAMEROVA, I., KUDRYASHOVA, E., VENKATRAMAN, G. & SPENCER, M. J. 2005. Calpain 3 participates in sarcomere remodeling by acting upstream of the ubiquitin-proteasome pathway. *Hum Mol Genet,* 14**,** 2125-34.

KREITMAN, R. J. 2006. Immunotoxins for targeted cancer therapy. *AAPS J,* 8**,** E532-51.

KRUTCHINSKY, A. N. & CHAIT, B. T. 2002. On the nature of the chemical noise in MALDI mass spectra. *J Am Soc Mass Spectrom,* 13**,** 129-34.

KUI WONG, N., EASTON, R. L., PANICO, M., SUTTON-SMITH, M., MORRISON, J. C., LATTANZIO, F. A., MORRIS, H. R., CLARK, G. F., DELL, A. & PATANKAR, M. S. 2003. Characterization of the oligosaccharides associated with the human ovarian tumor marker CA125. *J Biol Chem,* 278**,** 28619-34.

LAHIRI, S. & FUTERMAN, A. H. 2007. The metabolism and function of sphingolipids and glycosphingolipids. *Cell Mol Life Sci,* 64**,** 2270-84.

LANNERT, H., BUNNING, C., JECKEL, D. & WIELAND, F. T. 1994. Lactosylceramide is synthesized in the lumen of the Golgi apparatus. *FEBS Lett,* 342**,** 91-6.

LANNOO, N. & VAN DAMME, E. J. 2014. Lectin domains at the frontiers of plant defense. *Front Plant Sci,* 5**,** 397.

LARESGOITI-SERVITJE, E., GOMEZ-LOPEZ, N. & OLSON, D. M. 2010. An immunological insight into the origins of pre-eclampsia. *Hum Reprod Update,* 16**,** 510-24.

LAU, K. S., PARTRIDGE, E. A., GRIGORIAN, A., SILVESCU, C. I., REINHOLD, V. N., DEMETRIOU, M. & DENNIS, J. W. 2007. Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell,* 129**,** 123-34.

LEE, C. L., PANG, P. C., YEUNG, W. S., TISSOT, B., PANICO, M., LAO, T. T., CHU, I. K., LEE, K. F., CHUNG, M. K., LAM, K. K., KOISTINEN, R., KOISTINEN, H., SEPPALA, M., MORRIS, H. R., DELL, A. & CHIU, P. C. 2009. Effects of differential glycosylation of glycodelins on lymphocyte survival. *J Biol Chem,* 284**,** 15084-96.

LESSIN, D. L., HUNT, J. S., KING, C. R. & WOOD, G. W. 1988. Antigen expression by cells near the maternal-fetal interface. *Am J Reprod Immunol Microbiol,* 16**,** 1-7.

LIN, J. Y., TSERNG, K. Y., CHEN, C. C., LIN, L. T. & TUNG, T. C. 1970. Abrin and ricin: new anti-tumour substances. *Nature,* 227**,** 292-3.

LINGWOOD, C. A. 2011. Glycosphingolipid functions. *Cold Spring Harb Perspect Biol,* 3.

LIOLITSA, D., RAHMAN, S., BENTON, S., CARR, L. J. & HANNA, M. G. 2003. Is the mitochondrial complex I ND5 gene a hot-spot for MELAS causing mutations? *Ann Neurol,* 53**,** 128-32.

LIRA-NAVARRETE, E., DE LAS RIVAS, M., COMPANON, I., PALLARES, M. C., KONG, Y., IGLESIAS-FERNANDEZ, J., BERNARDES, G. J., PEREGRINA, J. M., ROVIRA, C., BERNADO, P., BRUSCOLINI, P., CLAUSEN, H., LOSTAO, A., CORZANA, F. & HURTADO-GUERRERO, R. 2015. Dynamic interplay between catalytic and lectin domains of GalNAc-transferases modulates protein O-glycosylation. *Nat Commun,* 6**,** 6937.

LIU, S., YANG, X., LIU, Y., WANG, X. & YAN, Q. 2011. sLeX/L-selectin mediates adhesion in vitro implantation model. *Mol Cell Biochem,* 350**,** 185-92.

LONLAY, P. & SETA, N. 2009. The clinical spectrum of phosphomannose isomerase deficiency, with an evaluation of mannose treatment for CDG-Ib. *Biochim Biophys Acta,* 1792**,** 841-3.

LOO, J. A., BROWN, J., CRITCHLEY, G., MITCHELL, C., ANDREWS, P. C. & OGORZALEK LOO, R. R. 1999. High sensitivity mass spectrometric methods for obtaining intact molecular weights from gel-separated proteins. *Electrophoresis,* 20**,** 743-8.

LORD, J. M. & ROBERTS, L. M. 1998. Toxin entry: retrograde transport through the secretory pathway. *J Cell Biol,* 140**,** 733-6.

LORD, J. M., ROBERTS, L. M. & ROBERTUS, J. D. 1994. Ricin: structure, mode of action, and some current applications. *FASEB J,* 8**,** 201-8.

LOWE, J. B. & MARTH, J. D. 2003. A genetic approach to Mammalian glycan function. *Annu Rev Biochem,* 72**,** 643-91.

LUCAS, H., BERCEGEAY, S., LE PENDU, J., JEAN, M., MIRALLIE, S. & BARRIERE, P. 1994. A fucose-containing epitope potentially involved in gamete interaction on the human zona pellucida. *Hum Reprod,* 9**,** 1532-8.

M NZENBERG, G. 2013. Development of mass spectrometers from Thomson and Aston to present. *Int J Mass Spectrom,* 349-350**,** 9-18.

MA, X., LIU, P., YAN, H., SUN, H., LIU, X., ZHOU, F., LI, L., CHEN, Y., MUTHANA, M. M., CHEN, X., WANG, P. G. & ZHANG, L. 2013. Substrate specificity provides insights into the sugar donor recognition mechanism of O-GlcNAc transferase (OGT). *PLoS One,* 8**,** e63452.

MAMYRIN, B. A., KARATAEV, V. I., SHMIKK, D. V. & ZAGULIN, V. A. 1973. Mass-Reflectron a New Nonmagnetic Time-of-Flight High-Resolution Mass-Spectrometer. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki,* 64**,** 82-89.

MARCUS, J., HONIGBAUM, S., SHROFF, S., HONKE, K., ROSENBLUTH, J. & DUPREE, J. L. 2006. Sulfatide is essential for the maintenance of CNS myelin and axon structure. *Glia,* 53**,** 372-81.

MART NEZ-DUNCKER, I., ASTEGGIANO, C. G. & FREEZE, H. H. 2012. Congenital disorders of glycosylation. *In:* MANUEL, H. & MONTES, M. (eds.) *Glycans: Biochemistry, Characterization and Applications.* Nova Science Pub Inc.

MARTIN, P. T. 2002. Glycobiology of the synapse. *Glycobiology,* 12**,** 1R-7R.

MASSOULIE, J. & BON, S. 1982. The molecular forms of cholinesterase and acetylcholinesterase in vertebrates. *Annu Rev Neurosci,* 5**,** 57-106.

MATALON, R., SURENDRAN, S., CAMPBELL, G. A., MICHALS-MATALON, K., TYRING, S. K., GRADY, J., CHENG, S. & KAYE, E. 2006. Hyaluronidase increases the biodistribution

of acid alpha-1,4 glucosidase in the muscle of Pompe disease mice: an approach to enhance the efficacy of enzyme replacement therapy. *Biochem Biophys Res Commun,* 350**,** 783-7.

MATSUI, T., TAKITA, E., SATO, T., KINJO, S., AIZAWA, M., SUGIURA, Y., HAMABATA, T., SAWADA, K. & KATO, K. 2011. N-glycosylation at noncanonical Asn-X-Cys sequences in plant cells. *Glycobiology,* 21**,** 994-9.

MATSUURA, A., ITO, M., SAKAIDANI, Y., KONDO, T., MURAKAMI, K., FURUKAWA, K., NADANO, D., MATSUDA, T. & OKAJIMA, T. 2008. O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors. *J Biol Chem,* 283**,** 35486-95.

MAVEYRAUD, L., NIWA, H., GUILLET, V., SVERGUN, D. I., KONAREV, P. V., PALMER, R. A., PEUMANS, W. J., ROUGE, P., VAN DAMME, E. J., REYNOLDS, C. D. & MOUREY, L. 2009. Structural basis for sugar recognition, including the Tn carcinoma antigen, by the lectin SNA-II from Sambucus nigra. *Proteins,* 75**,** 89-103.

MEDAWAR, P. B. Some immunological and endocrinological problems raised by the evolution of viviparity in vertebrates. Symp Soc Exp Biol, 1953. 38.

MERRILL, A. H., JR. 2002. De novo sphingolipid biosynthesis: a necessary, but dangerous, pathway. *J Biol Chem,* 277**,** 25843-6.

MERRILL, A. H., JR. 2011. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chem Rev,* 111**,** 6387-422.

MEYER, S., VAN LIEMPT, E., IMBERTY, A., VAN KOOYK, Y., GEYER, H., GEYER, R. & VAN DIE, I. 2005. DC-SIGN mediates binding of dendritic cells to authentic pseudo-LewisY glycolipids of Schistosoma mansoni cercariae, the first parasite-specific ligand of DC-SIGN. *J Biol Chem,* 280**,** 37349-59.

MONIATTE, M., VAN DER GOOT, F. G., BUCKLEY, J. T., PATTUS, F. & VAN DORSSELAER, A. 1996. Characterisation of the heptameric pore-forming complex of the Aeromonas toxin aerolysin using MALDI-TOF mass spectrometry. *FEBS Lett,* 384**,** 269-72.

MONK, C. R., SUTTON-SMITH, M., DELL, A. & GARDEN, O. A. 2006. Preparation of CD25(+) and CD25(-) CD4(+) T cells for glycomic analysis--a cautionary tale of serum glycoprotein sequestration. *Glycobiology,* 16**,** 11G-13G.

MONTFORT, W., VILLAFRANCA, J. E., MONZINGO, A. F., ERNST, S. R., KATZIN, B., RUTENBER, E., XUONG, N. H., HAMLIN, R. & ROBERTUS, J. D. 1987. The three-dimensional structure of ricin at 2.8 A. *J Biol Chem,* 262**,** 5398-403.

MORRIS, H. R., DELL, A., EASTON, R. L., PANICO, M., KOISTINEN, H., KOISTINEN, R., OEHNINGER, S., PATANKAR, M. S., SEPPALA, M. & CLARK, G. F. 1996a. Gender-specific glycosylation of human glycodelin affects its contraceptive activity. *Journal of Biological Chemistry,* 271**,** 32159-32167.

MORRIS, H. R., DELL, A. & MCDOWELL, R. A. 1981a. Extended performance using a high field magnet mass spectrometer. *Biomedical Mass Spectrometry,* 8**,** 463-73.

MORRIS, H. R., PANICO, M., BARBER, M., BORDOLI, R. S., SEDGWICK, R. D. & TYLER, A. 1981b. Fast atom bombardment: a new mass spectrometric method for peptide sequence analysis. *Biochem Biophys Res Commun,* 101**,** 623-31.

MORRIS, H. R., PAXTON, T., DELL, A., LANGHORNE, J., BERG, M., BORDOLI, R. S., HOYES, J. & BATEMAN, R. H. 1996b. High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun Mass Spectrom,* 10**,** 889-96.

MORRIS, H. R., PAXTON, T., PANICO, M., MCDOWELL, R. & DELL, A. 1997. A novel geometry mass spectrometer, the Q-TOF, for low-femtomole/attomole-range biopolymer sequencing. *J Protein Chem,* 16**,** 469-79.

MORRIS, H. R., THOMPSON, M. R., OSUGA, D. T., AHMED, A. I., CHAN, S. M., VANDENHEEDE, J. R. & FEENEY, R. E. 1978. Antifreeze glycoproteins from the blood of an antarctic fish. The structure of the proline-containing glycopeptides. *J Biol Chem,* 253**,** 5155-62.

MUNSON, M. S. B., FIELD, F. H. 1966. Chemical Ionization Mass Spectrometry. I. General Introduction. *J. Am. Chem. Soc.,* 88**,** 2621–2630.

MURPHY, S. P., CHOI, J. C. & HOLTZ, R. 2004. Regulation of major histocompatibility complex class II gene expression in trophoblast cells. *Reprod Biol Endocrinol,* 2**,** 52.

MURPHY, S. P. & TOMASI, T. B. 1998. Absence of MHC class II antigen expression in trophoblast cells results from a lack of class II transactivator (CIITA) gene expression. *Mol Reprod Dev,* 51**,** 1-12.

MYERS, S. A., DAOU, S., AFFAR, E. B. & BURLINGAME, A. 2013. Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments-a study of the O-GlcNAcylated protein Host Cell Factor C1. *Proteomics,* 13**,** 982-991.

NAKAISHI, Y., BANDO, M., SHIMIZU, H., WATANABE, K., GOTO, F., TSUGE, H., KONDO, K. & KOMATSU, M. 2009. Structural analysis of human glutamine:fructose-6-phosphate amidotransferase, a key regulator in type 2 diabetes. *Febs Letters,* 583**,** 163-167.

NARASIMHAN, S., FREED, J. C. & SCHACHTER, H. 1985. Control of Glycoprotein-Synthesis .10. Control of Glycoprotein-Synthesis - Bovine-Milk Udp-Galactose - N-Acetylglucosamine Beta-4-Galactosyltransferase Catalyzes the Preferential Transfer of Galactose to the Glcnac-Beta-1,2man-Alpha-1,3 Branch of Both Bisected and Nonbisected Complex Biantennary Asparagine-Linked Oligosaccharidess. *Biochemistry,* 24**,** 1694-1700.

NEILSON, K. A., ALI, N. A., MURALIDHARAN, S., MIRZAEI, M., MARIANI, M., ASSADOURIAN, G., LEE, A., VAN SLUYTER, S. C. & HAYNES, P. A. 2011. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics,* 11**,** 535-53.

NICOLL, G., NI, J., LIU, D., KLENERMAN, P., MUNDAY, J., DUBOCK, S., MATTEI, M. G. & CROCKER, P. R. 1999. Identification and characterization of a novel siglec, siglec-7, expressed by human natural killer cells and monocytes. *J Biol Chem,* 274**,** 34089-95.

NIER, A. O. 1955. Determination of Isotopic Masses and Abundances by Mass Spectrometry. *Science,* 121**,** 737-44.

NORTH, S. J., HITCHEN, P. G., HASLAM, S. M. & DELL, A. 2009. Mass spectrometry in the analysis of N-linked and O-linked glycans. *Current Opinion in Structural Biology,* 19**,** 498-506.

NORTH, S. J., JANG-LEE, J., HARRISON, R., CANIS, K., ISMAIL, M. N., TROLLOPE, A., ANTONOPOULOS, A., PANG, P. C., GRASSI, P., AL-CHALABI, S., ETIENNE, A. T., DELL, A. & HASLAM, S. M. 2010. Mass Spectrometric Analysis of Mutant Mice. *Methods in Enzymology, Vol 478: Glycomics,* 478**,** 27-77.

OATES, J. E., DELL, A., FUKUDA, M. & FUKUDA, M. N. 1985. A rapid mass-spectrometric procedure for probing the non-reducing structures of lactosaminoglycan-containing glycoconjugates. *Carbohydr Res,* 141**,** 149-52.

OGAWA, M., SAWAGUCHI, S., KAWAI, T., NADANO, D., MATSUDA, T., YAGI, H., KATO, K., FURUKAWA, K. & OKAJIMA, T. 2015. Impaired O-linked N-acetylglucosaminylation in the endoplasmic reticulum by mutated epidermal growth factor (EGF) domain-specific O-linked N-acetylglucosamine transferase found in Adams-Oliver syndrome. *J Biol Chem,* 290**,** 2137-49.

OHTA, Y., TSUKADA, Y. & SUGIMORI, T. 1989. Purification and properties of neuraminidase isozymes in Arthrobacter ureafaciens mutant. *J Biochem,* 106**,** 1086-9.

OKI, T., YAMAZAKI, K., KUROMITSU, J., OKADA, M. & TANAKA, I. 1999. cDNA cloning and mapping of a novel subtype of glutamine : fructose-6-phosphate amidotransferase (GFAT2) in human and mouse. *Genomics,* 57**,** 227-234.

OROZCO, A. F. & LEWIS, D. E. 2010. Flow cytometric analysis of circulating microparticles in plasma. *Cytometry A,* 77**,** 502-14.

PALACE, J. 2012. DOK7 congenital myasthenic syndrome. *Ann N Y Acad Sci,* 1275**,** 49-53.

PALACE, J., LASHLEY, D., NEWSOM-DAVIS, J., COSSINS, J., MAXWELL, S., KENNETT, R., JAYAWANT, S., YAMANASHI, Y. & BEESON, D. 2007. Clinical features of the DOK7 neuromuscular junction synaptopathy. *Brain,* 130**,** 1507-15.

PANG, P. C., CHIU, P. C. N., LEE, C. L., CHANG, L. Y., PANICO, M., MORRIS, H. R., HASLAM, S. M., KHOO, K. H., CLARK, G. F., YEUNG, W. S. B. & DELL, A. 2011. Human Sperm Binding Is Mediated by the Sialyl-Lewis(x) Oligosaccharide on the Zona Pellucida. *Science,* 333**,** 1761-1764.

PANG, P. C., TISSOT, B., DROBNIS, E. Z., SUTOVSKY, P., MORRIS, H. R., CLARK, G. F. & DELL, A. 2007. Expression of bisecting type and Lewisx/Lewisy terminated N-glycans on human sperm. *J Biol Chem,* 282**,** 36593-602.

PARC, A. L., KARAV, S., BELL, J. M., FRESE, S. A., LIU, Y., MILLS, D. A., BLOCK, D. E. & BARILE, D. 2015. A novel endo-beta-N-acetylglucosaminidase releases specific N-glycans depending on different reaction conditions. *Biotechnol Prog*.

PARHAM, P. 2004. NK cells and trophoblasts: partners in pregnancy. *J Exp Med,* 200**,** 951-5.

PARIA, B. C., REESE, J., DAS, S. K. & DEY, S. K. 2002. Deciphering the cross-talk of implantation: advances and challenges. *Science,* 296**,** 2185-8.

PARKER, W. D., JR. & PARKS, J. K. 2005. Mitochondrial ND5 mutations in idiopathic Parkinson's disease. *Biochem Biophys Res Commun,* 326**,** 667-9.

PARRY, S., LEDGER, V., TISSOT, B., HASLAM, S. M., SCOTT, J., MORRIS, H. R. & DELL, A. 2007. Integrated mass spectrometric strategy for characterizing the glycans from

glycosphingolipids and glycoproteins: direct identification of sialyl Le(x) in mice. *Glycobiology,* 17**,** 646-654.

PATANKAR, M. S., OZGUR, K., OEHNINGER, S., DELL, A., MORRIS, H., SEPPALA, M. & CLARK, G. F. 1997. Expression of glycans linked to natural killer cell inhibition on the human zona pellucida. *Mol Hum Reprod,* 3**,** 501-5.

PAUL, W. & STEINWEDEL, H. 1960. *Apparatus for separating charged particles of different specific charges*. United States patent application 2939952.

PETRUZZELLA, V., DI GIACINTO, G., SCACCO, S., PIEMONTE, F., TORRACO, A., CARROZZO, R., VERGARI, R., DIONISI-VICI, C., LONGO, D., TESSA, A., PAPA, S. & BERTINI, E. 2003. Atypical Leigh syndrome associated with the D393N mutation in the mitochondrial ND5 subunit. *Neurology,* 61**,** 1017-8.

PHILLIPS, M. L., NUDELMAN, E., GAETA, F. C., PEREZ, M., SINGHAL, A. K., HAKOMORI, S. & PAULSON, J. C. 1990. ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. *Science,* 250**,** 1130-2.

POLLEY, M. J., PHILLIPS, M. L., WAYNER, E., NUDELMAN, E., SINGHAL, A. K., HAKOMORI, S. & PAULSON, J. C. 1991. CD62 and endothelial cell-leukocyte adhesion molecule 1 (ELAM-1) recognize the same carbohydrate ligand, sialyl-Lewis x. *Proc Natl Acad Sci U S A,* 88**,** 6224-8.

PORTER, A. G. & JANICKE, R. U. 1999. Emerging roles of caspase-3 in apoptosis. *Cell Death Differ,* 6**,** 99-104.

POSTHUMUS, M. A., KISTEMAKER, P. G., MEUZELAAR, H. L. C. & TENNOEVERDEBRAUW, M. C. 1978. Laser Desorption-Mass Spectrometry of Polar Non-Volatile Bio-Organic Molecules. *Analytical Chemistry,* 50**,** 985-991.

PROHASKA, R., SCHENKELBRUNNER, H. & TUPPY, H. 1978. Enzymatic-Synthesis of Blood-Group Lewis-Specific Glycolipids. *European Journal of Biochemistry,* 84**,** 161-166.

QASBA, P. K., RAMAKRISHNAN, B. & BOEGGEMAN, E. 2008. Structure and Function of beta-1,4-Galactosyltransferase. *Current Drug Targets,* 9**,** 292-309.

RADEMACHER, T. W., PAREKH, R. B. & DWEK, R. A. 1988. Glycobiology. *Annu Rev Biochem,* 57**,** 785-838.

RADIONOVA, A., FILIPPOV, I. & DERRICK, P. J. 2015. In pursuit of resolution in time-of-flight mass spectrometry: A historical perspective. *Mass Spectrom Rev*.

RAPOPORT, T. A., JUNGNICKEL, B. & KUTAY, U. 1996. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu Rev Biochem,* 65**,** 271-303.

RAVAL, K. K., TAO, R., WHITE, B. E., DE LANGE, W. J., KOONCE, C. H., YU, J., KISHNANI, P. S., THOMSON, J. A., MOSHER, D. F., RALPHE, J. C. & KAMP, T. J. 2015. Pompe disease results in a Golgi-based glycosylation deficit in human induced pluripotent stem cell-derived cardiomyocytes. *J Biol Chem,* 290**,** 3121-36.

RICHARD, I., BROUX, O., ALLAMAND, V., FOUGEROUSSE, F., CHIANNILKULCHAI, N., BOURG, N., BRENGUIER, L., DEVAUD, C., PASTURAUD, P., ROUDAUT, C. & ET AL. 1995. Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell,* 81**,** 27-40.

RINEHART, K. L., JR. 1982. Fast atom bombardment mass spectrometry. *Science,* 218**,** 254-60.

ROBSON, A., HARRIS, L. K., INNES, B. A., LASH, G. E., ALJUNAIDY, M. M., APLIN, J. D., BAKER, P. N., ROBSON, S. C. & BULMER, J. N. 2012. Uterine natural killer cells initiate spiral artery remodeling in human pregnancy. *FASEB J,* 26**,** 4876-85.

ROCHA, C. T. & HOFFMAN, E. P. 2010. Limb–Girdle and Congenital Muscular Dystrophies: Current Diagnostics, Management, and Emerging Technologies. *Current neurology and neuroscience reports,* 10**,** 267-276.

ROSEMAN, S. 2001. Reflections on glycobiology. *Journal of Biological Chemistry,* 276**,** 41527-41542.

RUDGE, M. R. H. 1968. Theory of the Ionization of Atoms by Electron Impact. *Rev. Mod. Phys.,* 40**,** 564-590.

RUTENBER, E. & ROBERTUS, J. D. 1991. Structure of ricin B-chain at 2.5 A resolution. *Proteins,* 10**,** 260-9.

SANDVIG, K., OLSNES, S. & PIHL, A. 1978. Binding, uptake and degradation of the toxic proteins abrin and ricin by toxin-resistant cell variants. *Eur J Biochem,* 82**,** 13-23.

SANDVIG, K., PRYDZ, K., HANSEN, S. H. & VAN DEURS, B. 1991. Ricin transport in brefeldin A-treated cells: correlation between Golgi structure and toxic effect. *J Cell Biol,* 115**,** 971-81.

SANES, J. R. & CHENEY, J. M. 1982. Lectin binding reveals a synapse-specific carbohydrate in skeletal muscle. *Nature,* 300**,** 646-7.

SANES, J. R. & LICHTMAN, J. W. 1999. Development of the vertebrate neuromuscular junction. *Annu Rev Neurosci,* 22**,** 389-442.

SASSI, A., LAZAROSKI, S., WU, G., HASLAM, S. M., FLIEGAUF, M., MELLOULI, F., PATIROGLU, T., UNAL, E., OZDEMIR, M. A., JOUHADI, Z., KHADIR, K., BEN-KHEMIS, L., BEN-ALI, M., BEN-MUSTAPHA, I., BORCHANI, L., PFEIFER, D., JAKOB, T., KHEMIRI, M., ASPLUND, A. C., GUSTAFSSON, M. O., LUNDIN, K. E., FALK-SORQVIST, E., MOENS, L. N., GUNGOR, H. E., ENGELHARDT, K. R., DZIADZIO, M., STAUSS, H., FLECKENSTEIN, B., MEIER, R., PRAYITNO, K., MAUL-PAVICIC, A., SCHAFFER, S., RAKHMANOV, M., HENNEKE, P., KRAUS, H., EIBEL, H., KOLSCH, U., NADIFI, S., NILSSON, M., BEJAOUI, M., SCHAFFER, A. A., SMITH, C. I., DELL, A., BARBOUCHE, M. R. & GRIMBACHER, B. 2014. Hypomorphic homozygous mutations in phosphoglucomutase 3 (PGM3) impair immunity and increase serum IgE levels. *J Allergy Clin Immunol,* 133**,** 1410-9, 1419 e1-13.

SCHACHTER, H. 1991. The 'yellow brick road' to branched complex N-glycans. *Glycobiology,* 1**,** 453-61.

SCHROEDER, H. W., JR. & CAVACINI, L. 2010. Structure and function of immunoglobulins. *J Allergy Clin Immunol,* 125**,** S41-52.

SCHWARZ, F. & AEBI, M. 2011. Mechanisms and principles of N-linked protein glycosylation. *Curr Opin Struct Biol,* 21**,** 576-82.

SCUDDER, P., UEMURA, K., DOLBY, J., FUKUDA, M. N. & FEIZI, T. 1983. Isolation and Characterization of an Endo-Beta-Galactosidase from Bacteroides-Fragilis. *Biochemical Journal,* 213**,** 485-494.

SELCEN, D., SHEN, X. M., MILONE, M., BRENGMAN, J., OHNO, K., DEYMEER, F., FINKEL, R., ROWIN, J. & ENGEL, A. G. 2013. GFPT1-myasthenia: clinical, structural, and electrophysiologic heterogeneity. *Neurology,* 81**,** 370-8.

SENDEREK, J., MULLER, J. S., DUSL, M., STROM, T. M., GUERGUELTCHEVA, V., DIEPOLDER, I., LAVAL, S. H., MAXWELL, S., COSSINS, J., KRAUSE, S., MUELAS, N., VILCHEZ, J. J., COLOMER, J., MALLEBRERA, C. J., NASCIMENTO, A., NAFISSI, S., KARIMINEJAD, A., NILIPOUR, Y., BOZORGMEHR, B., NAJMABADI, H., RODOLICO, C., SIEB, J. P., STEINLEIN, O. K., SCHLOTTER, B., SCHOSER, B., KIRSCHNER, J., HERRMANN, R., VOIT, T., OLDFORS, A., LINDBERGH, C., URTIZBEREA, A., VON DER HAGEN, M., HUBNER, A., PALACE, J., BUSHBY, K., STRAUB, V., BEESON, D., ABICHT, A. & LOCHMULLER, H. 2011. Hexosamine Biosynthetic Pathway Mutations Cause Neuromuscular Transmission Defect. *American Journal of Human Genetics,* 88**,** 162-172.

SHAHIDI-NOGHABI, S., VAN DAMME, E. J., DE VOS, W. H. & SMAGGHE, G. 2011. Internalization of Sambucus nigra agglutinins I and II in insect midgut CF-203 cells. *Arch Insect Biochem Physiol,* 76**,** 211-22.

SHAHIDI-NOGHABI, S., VAN DAMME, E. J., IGA, M. & SMAGGHE, G. 2010. Exposure of insect midgut cells to Sambucus nigra L. agglutinins I and II causes cell death via caspase-dependent apoptosis. *J Insect Physiol,* 56**,** 1101-7.

SHANG, C., CHEN, Q., DELL, A., HASLAM, S. M., DE VOS, W. H. & VAN DAMME, E. J. 2015. The Cytotoxicity of Elderberry Ribosome-Inactivating Proteins Is Not Solely Determined by Their Protein Translation Inhibition Activity. *PLoS One,* 10**,** e0132389.

SHANG, C. J. & VAN DAMME, E. J. M. 2014. Comparative analysis of carbohydrate binding properties of Sambucus nigra lectins and ribosome-inactivating proteins. *Glycoconjugate Journal,* 31**,** 345-354.

SHIBUYA, N., GOLDSTEIN, I. J., BROEKAERT, W. F., NSIMBA-LUBAKI, M., PEETERS, B. & PEUMANS, W. J. 1987. The elderberry (Sambucus nigra L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence. *J Biol Chem,* 262**,** 1596-601.

SIGNOR, L. & ERBA, E. B. 2013. Matrix-assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) Mass Spectrometric Analysis of Intact Proteins Larger than 100 kDa. *Jove-Journal of Visualized Experiments.*

SILLENCE, D. J. 2007. New insights into glycosphingolipid functions--storage, lipid rafts, and translocators. *Int Rev Cytol,* 262**,** 151-89.

SOMERS, W. S., TANG, J., SHAW, G. D. & CAMPHAUSEN, R. T. 2000. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. *Cell,* 103**,** 467-79.

SORRELL, J. M. & CAPLAN, A. I. 2004. Fibroblast heterogeneity: more than skin deep. *J Cell Sci,* 117**,** 667-75.

SPIK, G., BAYARD, B., FOURNET, B., STRECKER, G., BOUQUELET, S. & MONTREUIL, J. 1975. Studies on glycoconjugates. LXIV. Complete structure of two carbohydrate units of human serotransferrin. *FEBS Lett,* 50**,** 296-9.

SPIRO, R. G. 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology,* 12**,** 43R-56R.

SPOONER, R. A. & LORD, J. M. 2015. Ricin trafficking in cells. *Toxins (Basel),* 7**,** 49-65.

STEENTOFT, C., VAKHRUSHEV, S. Y., JOSHI, H. J., KONG, Y., VESTER-CHRISTENSEN, M. B., SCHJOLDAGER, K. T., LAVRSEN, K., DABELSTEEN, S., PEDERSEN, N. B., MARCOS-SILVA, L., GUPTA, R., BENNETT, E. P., MANDEL, U., BRUNAK, S., WANDALL, H. H., LEVERY, S. B. & CLAUSEN, H. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J,* 32**,** 1478-88.

STEPHENS, W. E. 1946. A Pulsed Mass Spectrometer with Time Dispersion. *Bull. Am. Phys. Soc.,* 21.

STIRPE, F. 2004. Ribosome-inactivating proteins. *Toxicon,* 44**,** 371-83.

STIRPE, F. 2013. Ribosome-inactivating proteins: from toxins to useful proteins. *Toxicon,* 67**,** 12-6.

SUTTON-SMITH, M. & DELL, A. 2006. Analysis of Carbohydrates/Glycoproteins by Mass Spectrometry. *In:* CELIS, J. E. (ed.) *Cell Biology: A Laboratory Handbook.* 3rd ed.: Elsevier

SVENNERHOLM, L. 1963. Chromatographic Separation of Human Brain Gangliosides. *J Neurochem,* 10**,** 613-23.

TAKAHASHI, M., KUROKI, Y., OHTSUBO, K. & TANIGUCHI, N. 2009. Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. *Carbohydr Res,* 344**,** 1387-90.

TANAKA, K., WAKI, H., IDO, Y., AKITA, S., YOSHIDA, Y. & YOSHIDA, T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* 2**,** 151-153.

TANGVORANUNTAKUL, P., GAGNEUX, P., DIAZ, S., BARDOR, M., VARKI, N., VARKI, A. & MUCHMORE, E. 2003. Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc Natl Acad Sci U S A,* 100**,** 12045-50.

TARENTINO, A. L., GOMEZ, C. M. & PLUMMER, T. H., JR. 1985. Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Biochemistry,* 24**,** 4665-71.

TARRADE, A., LAI KUEN, R., MALASSINE, A., TRICOTTET, V., BLAIN, P., VIDAUD, M. & EVAIN-BRION, D. 2001. Characterization of human villous and extravillous trophoblasts isolated from first trimester placenta. *Lab Invest,* 81**,** 1199-211.

TASHIMA, Y. & STANLEY, P. 2014. Antibodies that detect O-linked beta-D-N-acetylglucosamine on the extracellular domain of cell surface glycoproteins. *J Biol Chem,* 289**,** 11132-42.

TAYLOR, M. E. & DRICKAMER, K. 2011. *Introduction to glycobiology,* Oxford ; New York, Oxford University Press.

TEINTENIER-LELIEVRE, M., JULIEN, S., JULIANT, S., GUERARDEL, Y., DUONOR-CERUTTI, M., DELANNOY, P. & HARDUIN-LEPERS, A. 2005. Molecular cloning and expression of a human hST8Sia VI (alpha2,8-sialyltransferase) responsible for the synthesis of the diSia motif on O-glycosylproteins. *Biochem J,* 392**,** 665-74.

TEJERO, J., JIMENEZ, P., QUINTO, E. J., CORDOBA-DIAZ, D., GARROSA, M., CORDOBA-DIAZ, M., GAYOSO, M. J. & GIRBES, T. 2015. Elderberries: a source of ribosome-inactivating proteins with lectin activity. *Molecules,* 20**,** 2364-87.

TEN HAGEN, K. G., FRITZ, T. A. & TABAK, L. A. 2003. All in the family: the UDP-GalNAc : polypeptide N-acetylgalactosaminyltransferases. *Glycobiology,* 13**,** 1R-16R.

THAN, N. G., BALOGH, A., ROMERO, R., KARPATI, E., EREZ, O., SZILAGYI, A., KOVALSZKY, I., SAMMAR, M., GIZURARSON, S., MATKO, J., ZAVODSZKY, P., PAPP, Z. & MEIRI, H. 2014. Placental Protein 13 (PP13) - A Placental Immunoregulatory Galectin Protecting Pregnancy. *Front Immunol,* 5**,** 348.

THAN, N. G., PICK, E., BELLYEI, S., SZIGETI, A., BURGER, O., BERENTE, Z., JANAKY, T., BORONKAI, A., KLIMAN, H., MEIRI, H., BOHN, H., THAN, G. N. & SUMEGI, B. 2004. Functional analyses of placental protein 13/galectin-13. *Eur J Biochem,* 271**,** 1065-78.

THAN, N. G., ROMERO, R., GOODMAN, M., WECKLE, A., XING, J., DONG, Z., XU, Y., TARQUINI, F., SZILAGYI, A., GAL, P., HOU, Z., TARCA, A. L., KIM, C. J., KIM, J. S., HAIDARIAN, S., UDDIN, M., BOHN, H., BENIRSCHKE, K., SANTOLAYA-FORGAS, J., GROSSMAN, L. I., EREZ, O., HASSAN, S. S., ZAVODSZKY, P., PAPP, Z. & WILDMAN, D. E. 2009. A primate subfamily of galectins expressed at the maternal-fetal interface that promote immune cell death. *Proc Natl Acad Sci U S A,* 106**,** 9731-6.

THAN, N. G., SUMEGI, B., THAN, G. N., BERENTE, Z. & BOHN, H. 1999. Isolation and sequence analysis of a cDNA encoding human placental tissue protein 13 (PP13), a new lysophospholipase, homologue of human eosinophil Charcot-Leyden Crystal protein. *Placenta,* 20**,** 703-10.

TISSOT, B., NORTH, S. J., CERONI, A., PANG, P. C., PANICO, M., ROSATI, F., CAPONE, A., HASLAM, S. M., DELL, A. & MORRIS, H. R. 2009. Glycoproteomics: Past, present and future. *Febs Letters,* 583**,** 1728-1735.

TRAVING, C. & SCHAUER, R. 1998. Structure, function and metabolism of sialic acids. *Cell Mol Life Sci,* 54**,** 1330-49.

TRETTER, V., ALTMANN, F. & MARZ, L. 1991. Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1----3 to the asparagine-linked N-acetylglucosamine residue. *Eur J Biochem,* 199**,** 647-52.

TURACA, L. T., DE FARIA, D. O., KYOSEN, S. O., TEIXEIRA, V. D., MOTTA, F. L., PESSOA, J. G., RODRIGUES, E. S. M., DE ALMEIDA, S. S., D'ALMEIDA, V., MUNOZ ROJAS, M. V., MARTINS, A. M. & PESQUERO, J. B. 2015. Novel GAA mutations in patients with Pompe disease. *Gene,* 561**,** 124-31.

UNIPROTKB. 2012. *GFPT1_HUMAN* [Online]. Available: http://www.uniprot.org/uniprot/Q06210.

VAN DAMME, E. J., BARRE, A., ROUGE, P., VAN LEUVEN, F. & PEUMANS, W. J. 1997a. Isolation and molecular cloning of a novel type 2 ribosome-inactivating protein with an inactive B chain from elderberry (Sambucus nigra) bark. *J Biol Chem,* 272**,** 8353-60.

VAN DAMME, E. J., HAO, Q., CHEN, Y., BARRE, A., VANDENBUSSCHE, F., DESMYTER, S., ROUG , P. & PEUMANS, W. J. 2001. Ribosome-Inactivating Proteins: A Family of Plant Proteins That Do More Than Inactivate Ribosomes. *Critical Reviews in Plant Sciences* 20**,** 395-465.

VAN DAMME, E. J., PEUMANS, W. J., PUSZTAI, A. & BARDOCZ, S. 1998. *Sambucus nigra. Handbook of Plant Lectins: Properties and Biomedical Applications.* Wiley.

VAN DAMME, E. J., ROY, S., BARRE, A., CITORES, L., MOSTAFAPOUS, K., ROUGE, P., VAN LEUVEN, F., GIRBES, T., GOLDSTEIN, I. J. & PEUMANS, W. J. 1997b. Elderberry (Sambucus nigra) bark contains two structurally different Neu5Ac(alpha2,6)Gal/GalNAc-binding type 2 ribosome-inactivating proteins. *Eur J Biochem,* 245**,** 648-55.

VAN DER PLOEG, A. T. & REUSER, A. J. 2008. Pompe's disease. *Lancet,* 372**,** 1342-53.

VARKI, A. 2007. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature,* 446**,** 1023-1029.

VARKI, A. 2009. *Essentials of glycobiology,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

VESTAL, M. L. & CAMPBELL, J. M. 2005. Tandem time-of-flight mass spectrometry. *Methods Enzymol,* 402**,** 79-108.

VESTAL, M. L., JUHASZ, P. & MARTIN, S. A. 1995. Delayed Extraction Matrix-Assisted Laser-Desorption Time-of-Flight Mass-Spectrometry. *Rapid Communications in Mass Spectrometry,* 9**,** 1044-1050.

VOET, D. & VOET, J. G. 2004. *Biochemistry,* New York, J. Wiley & Sons.

WADA, Y., AZADI, P., COSTELLO, C. E., DELL, A., DWEK, R. A., GEYER, H., GEYER, R., KAKEHI, K., KARLSSON, N. G., KATO, K., KAWASAKI, N., KHOO, K. H., KIM, S., KONDO, A., LATTOVA, E., MECHREF, Y., MIYOSHI, E., NAKAMURA, K., NARIMATSU, H., NOVOTNY, M. V., PACKER, N. H., PERREAULT, H., PETER-KATALINIC, J., POHLENTZ, G., REINHOLD, V. N., RUDD, P. M., SUZUKI, A. & TANIGUCHI, N. 2007. Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology,* 17**,** 411-22.

WADA, Y., DELL, A., HASLAM, S. M., TISSOT, B., CANIS, K., AZADI, P., BACKSTROM, M., COSTELLO, C. E., HANSSON, G. C., HIKI, Y., ISHIHARA, M., ITO, H., KAKEHI, K., KARLSSON, N., HAYES, C. E., KATO, K., KAWASAKI, N., KHOO, K. H., KOBAYASHI, K., KOLARICH, D., KONDO, A., LEBRILLA, C., NAKANO, M., NARIMATSU, H., NOVAK, J., NOVOTNY, M. V., OHNO, E., PACKER, N. H., PALAIMA, E., RENFROW, M. B., TAJIRI, M., THOMSSON, K. A., YAGI, H., YU, S. Y. & TANIGUCHI, N. 2010. Comparison of methods for profiling O-glycosylation: Human Proteome Organisation Human Disease Glycomics/Proteome Initiative multi-institutional study of IgA1. *Mol Cell Proteomics,* 9**,** 719-27.

WANG, X. Q., LEE, S., WILSON, H., SEEGER, M., IORDANOV, H., GATLA, N., WHITTINGTON, A., BACH, D., LU, J. Y. & PALLER, A. S. 2014. Ganglioside GM3 depletion reverses impaired wound healing in diabetic mice by activating IGF-1 and insulin receptors. *J Invest Dermatol,* 134**,** 1446-55.

WEERAPANA, E. & IMPERIALI, B. 2006. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology,* 16**,** 91R-101R.

WILLIAMS, T. L., ANDRZEJEWSKI, D., LAY, J. O. & MUSSER, S. M. 2003. Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. *Journal of the American Society for Mass Spectrometry,* 14**,** 342-351.

WITTE, K., SEARS, P., MARTIN, R. & WONG, C. H. 1997. Enzymatic glycoprotein synthesis: Preparation of ribonuclease glycoforms via enzymatic glycopeptide condensation and glycosylation. *Journal of the American Chemical Society,* 119**,** 2114-2118.

WITTMANN, C. 2007. Fluxome analysis using GC-MS. *Microbial Cell Factories,* 6.

WONG, T., MCGRATH, J. A. & NAVSARIA, H. 2007. The role of fibroblasts in tissue engineering and regeneration. *Br J Dermatol,* 156**,** 1149-55.

WOPEREIS, S., LEFEBER, D. J., MORAVA, E. & WEVERS, R. A. 2006. Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: a review. *Clin Chem,* 52**,** 574-600.

YAMAJI, T., TERANISHI, T., ALPHEY, M. S., CROCKER, P. R. & HASHIMOTO, Y. 2002. A small region of the natural killer cell receptor, Siglec-7, is responsible for its preferred binding to alpha 2,8-disialyl and branched alpha 2,6-sialyl residues. A comparison with Siglec-9. *J Biol Chem,* 277**,** 6324-32.

YAMASHITA, T., WADA, R., SASAKI, T., DENG, C., BIERFREUND, U., SANDHOFF, K. & PROIA, R. L. 1999. A vital role for glycosphingolipid synthesis during development and differentiation. *Proc Natl Acad Sci U S A,* 96**,** 9142-7.

YANG, R. Y., RABINOVICH, G. A. & LIU, F. T. 2008. Galectins: structure, function and therapeutic potential. *Expert Rev Mol Med,* 10**,** e17.

YOSHIDA-MORIGUCHI, T., YU, L. P., STALNAKER, S. H., DAVIS, S., KUNZ, S., MADSON, M., OLDSTONE, M. B. A., SCHACHTER, H., WELLS, L. & CAMPBELL, K. P. 2010. O-Mannosyl Phosphorylation of Alpha-Dystroglycan Is Required for Laminin Binding. *Science,* 327**,** 88-92.

YOSHIMURA, M., IHARA, Y., OHNISHI, A., IJUHIN, N., NISHIURA, T., KANAKURA, Y., MATSUZAWA, Y. & TANIGUCHI, N. 1996. Bisecting N-acetylglucosamine on K562 cells suppresses natural killer cytotoxicity and promotes spleen colonization. *Cancer Res,* 56**,** 412-8.

YOSHIMURA, Y., NUDELMAN, A. S., LEVERY, S. B., WANDALL, H. H., BENNETT, E. P., HINDSGAUL, O., CLAUSEN, H. & NISHIMURA, S. 2012. Elucidation of the sugar recognition ability of the lectin domain of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase 3 by using unnatural glycopeptide substrates. *Glycobiology,* 22**,** 429-38.

YU, S. Y., WU, S. W. & KHOO, K. H. 2006. Distinctive characteristics of MALDI-Q/TOF and TOF/TOF tandem mass spectrometry for sequencing of permethylated complex type N-glycans. *Glycoconjugate Journal,* 23**,** 355-369.

ZAIA, J. 2004. Mass spectrometry of oligosaccharides. *Mass Spectrom Rev,* 23**,** 161-227.

ZENOBI, R. & KNOCHENMUSS, R. 1998. Ion formation in MALDI mass spectrometry. *Mass Spectrometry Reviews,* 17**,** 337-366.

ZHANG, M., WANG, M., GAO, R., LIU, X., CHEN, X., GENG, Y., DING, Y., WANG, Y. & HE, J. 2015. Altered beta1,6-GlcNAc and bisecting GlcNAc-branched N-glycan on integrin beta1 are associated with early spontaneous miscarriage in humans. *Hum Reprod*.

ZHAO, Y., SATO, Y., ISAJI, T., FUKUDA, T., MATSUMOTO, A., MIYOSHI, E., GU, J. & TANIGUCHI, N. 2008. Branched N-glycans regulate the biological functions of integrins and cadherins. *FEBS J,* 275**,** 1939-48.

ZOLTOWSKA, K., WEBSTER, R., FINLAYSON, S., MAXWELL, S., COSSINS, J., MULLER, J., LOCHMULLER, H. & BEESON, D. 2013. Mutations in GFPT1 that underlie limb-girdle congenital myasthenic syndrome result in reduced cell-surface expression of muscle AChR. *Human Molecular Genetics,* 22**,** 2905-2913.